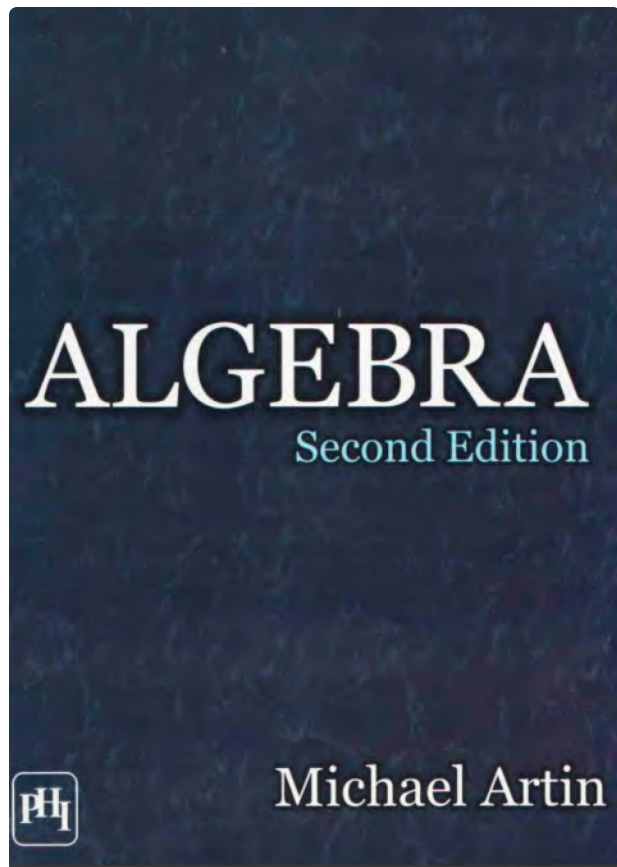


# Algebra 2ed by Michael Artin

---



- Complete solution guide to almost every exercise of Algebra by Michael Artin 2ed from quizlet & chegg

# 1

## Chapter 1

### Section 1

1. a

Comparing

$$\begin{bmatrix} 1 & 2 & 5 \\ 2 & 7 & 8 \\ 0 & 9 & 4 \end{bmatrix} \text{ with } \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

we get  $a_{21} = 2$  and  $a_{23} = 8$ .

**Result**

$$a_{21} = 2 \text{ and } a_{23} = 8.$$

2. a

For given

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -8 & -4 \\ 9 & 5 \\ -3 & -2 \end{bmatrix}$$

$AB$  is given by

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} -8 & -4 \\ 9 & 5 \\ -3 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \cdot (-8) + 2 \cdot 9 + 3 \cdot (-3) & 1 \cdot (-4) + 2 \cdot 5 + 3 \cdot (-2) \\ 3 \cdot (-8) + 3 \cdot 9 + 1 \cdot (-3) & 3 \cdot (-4) + 3 \cdot 5 + 1 \cdot (-2) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

For given  $A$  and  $B$ ,  $BA$  is given by

$$\begin{aligned} BA &= \begin{bmatrix} -8 & -4 \\ 9 & 5 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} (-8) \cdot 1 + (-4) \cdot 3 & (-8) \cdot 2 + (-4) \cdot 3 & (-8) \cdot 3 + (-4) \cdot 1 \\ 9 \cdot 1 + 5 \cdot 3 & 9 \cdot 2 + 5 \cdot 3 & 9 \cdot 3 + 5 \cdot 1 \\ (-3) \cdot 1 + (-2) \cdot 3 & (-3) \cdot 2 + (-2) \cdot 3 & (-3) \cdot 3 + (-2) \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} -20 & -28 & -28 \\ 24 & 33 & 32 \\ -9 & -12 & -11 \end{bmatrix} \end{aligned}$$



For given

$$A = \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & -4 \\ 3 & 2 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 6 & -4 \\ 3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \cdot 6 + 4 \cdot 3 & 1 \cdot (-4) + 4 \cdot 2 \\ 1 \cdot 6 + 2 \cdot 3 & 1 \cdot (-4) + 4 \cdot 2 \end{bmatrix}$$

$$= \begin{bmatrix} 18 & 4 \\ 12 & 0 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} 6 & -4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6 \cdot 1 + (-4) \cdot 1 & 6 \cdot 4 + (-4) \cdot 2 \\ 3 \cdot 1 + 2 \cdot 1 & 3 \cdot 4 + 2 \cdot 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 16 \\ 5 & 16 \end{bmatrix}$$

**Result**

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$BA = \begin{bmatrix} -20 & -28 & -28 \\ 24 & 33 & 32 \\ -9 & -12 & -11 \end{bmatrix}$$

$$AB = \begin{bmatrix} 18 & 4 \\ 12 & 0 \end{bmatrix}, BA = \begin{bmatrix} 2 & 16 \\ 5 & 16 \end{bmatrix}$$

3. a

Given  $A = [a_1 \ a_2 \ \dots \ a_n]$  a row vector and

$$B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

a column vector.

$$AB = [a_1 \ a_2 \ \dots \ a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$= [a_1 b_1 + a_2 b_2 + \dots + a_n b_n]$$

which is  $1 \times 1$  matrix.

$$BA = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} [a_1 \ a_2 \ \dots \ a_n]$$

$$= \begin{bmatrix} b_1 a_1 & b_1 a_2 & \dots & b_1 a_n \\ b_2 a_1 & b_2 a_2 & \dots & b_2 a_n \\ \vdots & \vdots & \dots & \vdots \\ b_n a_1 & b_n a_2 & \dots & b_n a_n \end{bmatrix}$$

**Result**

$AB = [a_1 b_1 + a_2 b_2 + \dots + a_n b_n]$  and

$$BA = \begin{bmatrix} b_1 a_1 & b_1 a_2 & \dots & b_1 a_n \\ b_2 a_1 & b_2 a_2 & \dots & b_2 a_n \\ \vdots & \vdots & \dots & \vdots \\ b_n a_1 & b_n a_2 & \dots & b_n a_n \end{bmatrix}$$

#### 4. a

We have

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 \cdot 0 + 2 \cdot 1 & 1 \cdot 1 + 2 \cdot 1 & 1 \cdot 2 + 2 \cdot 3 \\ 0 \cdot 0 + 1 \cdot 1 & 0 \cdot 1 + 1 \cdot 1 & 0 \cdot 2 + 1 \cdot 3 \end{bmatrix} \\ = \begin{bmatrix} 2 & 3 & 8 \\ 1 & 1 & 3 \end{bmatrix}$$

Now,

$$\left( \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 8 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} \\ = \begin{bmatrix} 2 \cdot 1 + 3 \cdot 4 + 8 \cdot 3 \\ 1 \cdot 1 + 1 \cdot 4 + 3 \cdot 3 \end{bmatrix} \\ = \begin{bmatrix} 38 \\ 14 \end{bmatrix}$$

Now,

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + 1 \cdot 4 + 2 \cdot 3 \\ 1 \cdot 1 + 1 \cdot 4 + 3 \cdot 3 \end{bmatrix} \\ = \begin{bmatrix} 10 \\ 14 \end{bmatrix}$$

Now,

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \left( \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 14 \end{bmatrix} \\ = \begin{bmatrix} 1 \cdot 10 + 2 \cdot 14 \\ 0 \cdot 10 + 1 \cdot 14 \end{bmatrix} \\ = \begin{bmatrix} 38 \\ 14 \end{bmatrix}$$

This shows that matrix product is associative.

#### Result

Show that

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \left( \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} \right) = \left( \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$$

#### 5. a

Given  $A$ ,  $B$  and  $C$  are  $l \times m$ ,  $m \times n$  and  $n \times p$ . In computation of  $AB$ , in each cell, row of  $A$  multiplies column of  $B$  for which there are  $m$  multiplications. Since there are  $l \times n$  elements in  $AB$ , there are total of  $l \times m \times n$  multiplications.

Computing  $ABC$  as  $(AB)C$ , there are  $l \times m \times n + l \times n \times p$  multiplications. Computing  $ABC$  as  $A(BC)$ , there are  $m \times n \times p + l \times m \times p$ . To minimize computation, we may choose order  $A(BC)$  if  $l \times m \times n + l \times n \times p$  is smaller than  $m \times n \times p + l \times m \times p$ . Otherwise we may choose  $(AB)C$ .

#### Result

2 of 2

There are  $l \times m \times n$  multiplication in  $AB$ . If  $l$  is large then we may choose  $A(BC)$ . If  $p$  is large then we may choose  $(AB)C$ .

6. a

We have

$$\begin{bmatrix} 1 & a \\ & 1 \end{bmatrix} \begin{bmatrix} 1 & b \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & a+b \\ & 1 \end{bmatrix}$$

This with  $b = a$ , this gives

$$\begin{aligned} \begin{bmatrix} 1 & a \\ & 1 \end{bmatrix}^2 &= \begin{bmatrix} 1 & 2a \\ & 1 \end{bmatrix} \\ \begin{bmatrix} 1 & a \\ & 1 \end{bmatrix}^3 &= \begin{bmatrix} 1 & 3a \\ & 1 \end{bmatrix} \\ &\vdots \\ \begin{bmatrix} 1 & a \\ & 1 \end{bmatrix}^n &= \begin{bmatrix} 1 & na \\ & 1 \end{bmatrix} \end{aligned}$$

### Result

$$\begin{bmatrix} 1 & a \\ & 1 \end{bmatrix} \begin{bmatrix} 1 & b \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & a+b \\ & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & a \\ & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & na \\ & 1 \end{bmatrix}$$

7. a

We calculate as follows

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^2 &= \begin{bmatrix} 1 & 2 & 3 \\ & 0 & 1 & 2 \\ & & 0 & 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^3 &= \begin{bmatrix} 1 & 3 & 6 \\ & 0 & 1 & 3 \\ & & 0 & 0 & 1 \end{bmatrix} \\ &\vdots \\ \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^n &= \begin{bmatrix} 1 & n & \frac{1}{2}n(n+1) \\ & 0 & 1 & n \\ & & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

To prove this inductively, for  $n = 1$ , this is trivially true. Suppose it were to hold for  $n$ . i.e.

$$\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n & \frac{1}{2}n(n+1) \\ & 0 & 1 & n \\ & & 0 & 0 & 1 \end{bmatrix}$$

Then

$$\begin{aligned}
 \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^{n+1} &= \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^n \\
 &= \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & n & \frac{1}{2}n(n+1) \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & n+1 & \frac{1}{2}n(n+1) + n+1 \\ & 1 & n+1 \\ & & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & n+1 & \frac{1}{2}(n+1)(n+2) \\ & 1 & n+1 \\ & & 1 \end{bmatrix}
 \end{aligned}$$

Hence, by mathematical induction, this holds for all all natural numbers  $n$ .

### Result

$$\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n & \frac{1}{2}n(n+1) \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$$

8. a

To calculate matrix, we may operate blockwise

$$\left[ \begin{array}{cc|cc} 1 & 1 & 1 & 5 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array} \right] \left[ \begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 3 \end{array} \right]$$

The top left component is given by

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix}$$

The top right component is given by

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 17 \\ 1 & 4 \end{bmatrix}$$

Bottom left component is given by

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}$$

Bottom right component is given by

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$$

Thus,

$$\left[ \begin{array}{cccc} 1 & 1 & 1 & 5 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array} \right] \left[ \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 3 \end{array} \right] = \left[ \begin{array}{cccc} 2 & 8 & 6 & 17 \\ 0 & 2 & 1 & 4 \\ 1 & 3 & 2 & 3 \\ 1 & 1 & 0 & 2 \end{array} \right]$$

Similarly,

$$\left[ \begin{array}{c|cc} 0 & 1 & 2 \\ \hline 0 & 1 & 0 \\ 3 & 0 & 1 \end{array} \right] \left[ \begin{array}{c|ccc} 1 & 2 & 3 \\ \hline 4 & 2 & 3 \\ 5 & 0 & 4 \end{array} \right]$$

Top left component is given by

$$[0][1] + [1 \ 2] \begin{bmatrix} 4 \\ 5 \end{bmatrix} = [14]$$

Top right component is given by

$$[0] \begin{bmatrix} 2 & 3 \end{bmatrix} + [1 \ 2] \begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix} = [2 \ 11]$$

Bottom left component is given by

$$\begin{bmatrix} 0 \\ 3 \end{bmatrix} [1] + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

Bottom right component is given by

$$\begin{bmatrix} 0 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 6 & 13 \end{bmatrix}$$

Thus,

$$\left[ \begin{array}{ccc} 0 & 1 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{array} \right] \left[ \begin{array}{ccc} 1 & 2 & 3 \\ 4 & 2 & 3 \\ 5 & 0 & 4 \end{array} \right] = \left[ \begin{array}{ccc} 14 & 2 & 11 \\ 4 & 2 & 3 \\ 8 & 6 & 13 \end{array} \right]$$

## Result

$$\left[ \begin{array}{cccc} 1 & 1 & 1 & 5 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array} \right] \left[ \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 3 \end{array} \right] = \left[ \begin{array}{cccc} 2 & 8 & 6 & 17 \\ 0 & 2 & 1 & 4 \\ 1 & 3 & 2 & 3 \\ 1 & 1 & 0 & 2 \end{array} \right]$$

$$\left[ \begin{array}{ccc} 0 & 1 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{array} \right] \left[ \begin{array}{ccc} 1 & 2 & 3 \\ 4 & 2 & 3 \\ 5 & 0 & 4 \end{array} \right] = \left[ \begin{array}{ccc} 14 & 2 & 11 \\ 4 & 2 & 3 \\ 8 & 6 & 13 \end{array} \right]$$

9. a

(a) Given  $A$  and  $B$  are square matrices. Using distributive property of matrices

$$\begin{aligned}(A + B)(A - B) &= A(A - B) + B(A - B) \\ &= A^2 - AB + BA - B^2\end{aligned}$$

Now if  $AB = BA$  i.e. the multiplication of two matrices commute, then

$$(A - B)(A + B) = A^2 - B^2$$

(b) Now again, using distributive property of matrix multiplication we get

$$\begin{aligned}(A + B)^3 &= (A + B)(A + B)(A + B) \\ &= (A + B)(AA + AB + BA + BB) \\ &= (A + B)(A^2 + AB + BA + B^2) \\ &= A^3 + A^2B + ABA + AB^2 + BA^2 + BAB + B^2A + B^3\end{aligned}$$

## Result

2 of 2

(a) If  $AB = BA$  then  $(A + B)(A - B) = A^2 - B^2$ . (b)  $(A + B)^3 = A^3 + A^2B + ABA + AB^2 + BA^2 + BAB + B^2A + B^3$

10. a

Given  $D$  be the diagonal matrix with diagonal entries  $d_1, \dots, d_n$ , and let  $A = (a_{ij})$  be an arbitrary  $n \times n$  matrix.

$$\begin{aligned}AD &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & & d_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11}d_1 & a_{12}d_2 & \cdots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \cdots & a_{2n}d_n \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nn}d_n \end{bmatrix} \\ &= (a_{ij}d_j)\end{aligned}$$

Similarly,

$$\begin{aligned}DA &= \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & & d_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix} \\ &= \begin{bmatrix} d_1a_{11} & d_1a_{12} & \cdots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \cdots & d_2a_{2n} \\ \vdots & & & \\ d_na_{n1} & d_na_{n2} & \cdots & d_na_{nn} \end{bmatrix} \\ &= (d_ia_{ij})\end{aligned}$$

## Result

$$AD = a_{ij}d_j \text{ and } DA = (d_ia_{ij})$$

11. a

Let  $A = (a_{ij}), B = (b_{ij})$  be two  $n \times n$  upper triangular matrices. i.e.  $a_{ij} = b_{ij} = 0$  for all  $j \leq i$ . Suppose  $C = AB = (c_{ij})$  then

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + \cdots + a_{in}b_{nj} \\ &= \sum_{k=1}^n a_{ik}b_{kj} \end{aligned}$$

Since  $A$  and  $B$  are upper triangular matrices, the sum becomes

$$\sum_{k=1}^n a_{ik}b_{kj} = \sum_{k=i+1}^n a_{ik}b_{kj}$$

If  $j \leq i$  then  $b_{kj} = 0$  since  $k \geq i+1 \geq j$ . This shows that

$$C = AB = (c_{ij}), \quad c_{ij} = 0 \quad \text{if} \quad i \geq j$$

Thus  $AB = C$  is upper triangular matrix.

## Result

2 of 2

Show  $AB = (c_{ij})$  where  $c_{ij} = 0$  if  $i \geq j$ .

12. a

(a) We have

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \end{aligned}$$

So if  $b = c = 0$  i.e. all matrices of the form

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

commutes with

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

(b) We have

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \\ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix} \end{aligned}$$

This gives  $c = 0, a = d$ . So all matrices of the form

$$\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$$

commutes with

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

(c) We have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 2a & 6b \\ 2c & 6d \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2a & 2b \\ 6c & 6d \end{bmatrix}$$

This gives  $b = c = 0$ . So all matrices of the form

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

commutes with

$$\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$$

(d) We have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 2a & 3a + 6b \\ 2c & 3c + 6d \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2a + 3c & 2b + 3d \\ 6c & 6d \end{bmatrix}$$

This

gives  $c = 0, 3a + 6b = 2b + 3d \implies d = a + \frac{4}{3}b$ .

So all matrices of the form

$$\begin{bmatrix} a & b \\ 0 & a + \frac{4}{3}b \end{bmatrix}$$

commutes with

$$\begin{bmatrix} 2 & 3 \\ 0 & 6 \end{bmatrix}$$

(a)

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

(b)

$$\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$$

(c)

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

(d)

$$\begin{bmatrix} a & b \\ 0 & a + \frac{4}{3}b \end{bmatrix}$$

13. a

Consider  $I - A^k$ , factoring it gives

$$I + A^k = (I + A)(I - A + A^2 - \dots + (-1)^{k-1}A^{k-1})$$

Since  $A^k = 0$ , we get

$$(I + A)(I - A + A^2 - \dots + (-1)^{k-1}A^{k-1}) = I$$

Let  $B = I - A + A^2 - \dots + (-1)^{k-1}A^{k-1}$ , then

$$B(I + A) = I + A - A + \dots + (-1)^{k-1}A^{k-1} - (-1)^kA^k = I$$

This shows  $B$  is inverse of  $I + A$  hence it is invertible.

## Result

Factorize  $I^k + A^k$  to get inverse of  $I + A$ .



14. a

Given,

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$$

which is  $3 \times 2$  matrix so if  $BA = I_2$  then  $B$  must be  $2 \times 3$  matrix. Let

$$B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

Then

$$BA = \begin{bmatrix} 2a+b+c & 3a+2b+c \\ 2d+e+f & 3d+2e+f \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Comparing gives us

$$\begin{aligned} 2a+b+c &= 1 \\ 3a+2b+c &= 0 \\ 2d+e+f &= 0 \\ 3d+2e+f &= 1 \end{aligned}$$

The system of linear equations described by above is consistent and has infinitely many solutions. Solving it gives us  $b = -a - 1, c = 2 - a, e = 1 - d, f = -d - 1$ . So for any value of  $a$  and  $d$ ,

$$\begin{bmatrix} a & -a-1 & 2-a \\ d & 1-d & -d-1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Basically the reason why there is no matrix  $C$  such that  $AC = I_3$  is  $C$  must be  $3 \times 2$  matrix which has 6 variables.  $AC = I_3$  is  $3 \times 3$  matrix which yields 9 equations. 9 equations on 6 variables cannot be consistent unless at least 3 equations are linear combination of the other 6. Suppose

$$C = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

Then

$$AC = \begin{bmatrix} 2a+3d & 2b+3e & 2c+3f \\ a+2d & b+2e & c+2f \\ a+d & b+e & c+f \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, we get

$$\begin{aligned} 2a+3d &= 1 \\ 2b+3e &= 0 \\ 2c+3f &= 0 \\ a+2d &= 0 \\ b+2e &= 1 \\ c+2f &= 0 \\ a+d &= 0 \\ b+e &= 0 \\ c+f &= 1 \end{aligned}$$

Here, we have equations  $b+e=0, b+2e=1$  and  $2b+3e=0$  which does not have solutions. Thus no such  $C$  exists.

## Result

Show that

$$\begin{bmatrix} a & -a-1 & 2-a \\ d & 1-d & -d-1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and  $AC = I_3$  yields 9 equations in 6 variables which is inconsistent.

15. a

Let  $A = (a_{ij})$  be arbitrary matrix, then

$$A = \sum_{k,l} e_{kl} a_{kl}$$

Now, using 1.1.23

$$e_{ij}A = \sum_{k,l} e_{ij}e_{kl}a_{kl} = \sum_l e_{il}a_{jl}$$

$$Ae_{ij} = \sum_{k,l} e_{kl}e_{ij}a_{kl} = \sum_k e_{kj}a_{ki}$$

Now, using  $e_j = \sum_k e_{kj}$

$$\begin{aligned} e_j A e_k &= e_j \sum_{l,m} e_{lm} a_{lm} \sum_{\eta} e_{\eta k} \\ &= e_j \sum_{l,m} e_{lk} a_{lm} \\ &= e_j \sum_l e_{lk} \sum_m a_{lm} \\ &= \sum_{\eta} e_{\eta j} \sum_l e_{lk} \sum_m a_{lm} \\ &= \sum_{\eta,l} e_{\eta j} e_{lk} \sum_m a_{lm} \\ &= \sum_{\eta} e_{\eta k} \sum_m a_{jm} \end{aligned}$$

For  $e_{ii}Ae_{jj}$ ,

$$\begin{aligned} e_{ii}Ae_{jj} &= \sum_{l,m} e_{ii}e_{lm}a_{lm}e_{jj} \\ &= \sum_m e_{im}e_{jj}a_{im} \\ &= e_{ij}a_{ij} \end{aligned}$$

Now, finally

$$\begin{aligned} e_{ij}Ae_{kl} &= \sum_{\eta,m} e_{ij}e_{\eta m}a_{\eta m}e_{kl} \\ &= \sum_m e_{im}a_{jm}e_{kl} \\ &= a_{jk}e_{il} \end{aligned}$$

## Result

$$\begin{aligned} e_{ij}A &= \sum_l e_{il}a_{jl}, & Ae_{ij} &= \sum_k e_{kj}a_{ki} \\ e_j A e_k &= \sum_{\eta} e_{\eta k} \sum_m a_{jm}, & e_{ii}Ae_{jj} &= e_{ij}a_{ij} \\ e_{ij}Ae_{kl} &= a_{jk}e_{il} \end{aligned}$$

## Section 2

1. a

Given

$$M = \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 1 & 1 & 2 & 6 & 10 \\ 1 & 2 & 5 & 2 & 7 \end{bmatrix} \rightarrow \rightarrow \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 3 & 1 & 2 \end{bmatrix}$$

Here, first row is subtracted from second and third row. The corresponding matrix operator is given by

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

The second operation is second and third row are exchanged.

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 3 & 1 & 2 \end{bmatrix} \rightarrow \rightarrow \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 5 & 5 \end{bmatrix}$$

The corresponding operation is given by

$$A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The third operation

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 5 & 5 \end{bmatrix} \rightarrow \rightarrow \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

in which second row is subtracted from first row. The corresponding operation is given by

$$A_3 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Our final matrix is

$$\begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

. Now, we must verify

$$A_3 A_2 A_1 M = \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

. We compute

$$A_3 A_2 A_1 = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

and

$$A_3 A_2 A_1 M = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 1 & 1 & 2 & 6 & 10 \\ 1 & 2 & 5 & 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 5 & 5 \end{bmatrix}$$

which is same as the matrix we obtained by row operation.

## Result

Compute

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

,

$$A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

and

$$A_3 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

. Then show that

$$A_3 A_2 A_1 M = \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 5 & 5 \end{bmatrix}$$

2. a

Given

$$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 0 & 0 & 4 \\ 1 & -4 & -2 & 2 \end{bmatrix}$$

(a) For

$$B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

, Applying  $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - R_1$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 3 & 0 & 0 & 4 & 0 \\ 1 & -4 & -2 & 2 & 0 \end{array} \right] \rightarrow \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \end{array} \right]$$

Again applying  $R_3 \rightarrow R_3 - R_2$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \end{array} \right] \rightarrow \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Applying  $R_2 \rightarrow -\frac{R_2}{6}$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Applying  $R_1 \rightarrow R_1 - R_2$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \rightarrow \left[ \begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ 1 & 0 & 0 & \frac{4}{3} & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

This gives us free to choose  $x_3$  and  $x_4$ , so set  $x_3 = s$  and  $x_4 = t$ .

$$x_1 = -\frac{4}{3}t, x_2 = -\frac{1}{2}s + \frac{1}{6}t$$

Therefore

$$X = \begin{bmatrix} -\frac{4}{3}t \\ -\frac{1}{2}s + \frac{1}{6}t \\ s \\ t \end{bmatrix}$$

(b) For

$$B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

, Applying  $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - R_1$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 3 & 0 & 0 & 4 & 1 \\ 1 & -4 & -2 & 2 & 0 \end{array} \right] \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -6 & -3 & 1 & -2 \\ 0 & -6 & -3 & 1 & -1 \end{array} \right]$$

Again applying  $R_3 \rightarrow R_3 - R_2$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -6 & -3 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

We note that on third row, we get  $0x_1 + 0x_2 + 0x_3 + 0x_4 = 1$  which is false. Therefore the system is inconsistent hence no solution exists.

(c) For

$$B = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

, Applying  $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - R_1$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 3 & 0 & 0 & 4 & 2 \\ 1 & -4 & -2 & 2 & 2 \end{array} \right] \rightarrow \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 2 \\ 0 & -6 & -3 & 1 & 2 \end{array} \right]$$

Again applying  $R_3 \rightarrow R_3 - R_2$ , we get

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 2 \\ 0 & -6 & -3 & 1 & 2 \end{array} \right] \rightarrow \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Applying  $R_2 \rightarrow -\frac{R_2}{6}$ ,

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \rightarrow \rightarrow \left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Applying  $R_1 \rightarrow R_1 - R_2$ ,

$$\left[ \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \rightarrow \rightarrow \left[ \begin{array}{cccc|c} 1 & 0 & 0 & \frac{4}{3} & \frac{2}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

This gives us free to choose  $x_3$  and  $x_4$ . Set  $x_3 = s, x_4 = t$  then we get

$$x_1 + \frac{4}{3}t = \frac{2}{3}, x_2 + \frac{1}{2}s - \frac{1}{6}t = -\frac{1}{3}$$

Hence solution is of the form

$$X = \begin{bmatrix} \frac{2}{3} - \frac{4}{3}t \\ -\frac{1}{3} - \frac{1}{2}s + \frac{1}{6}t \\ s \\ t \end{bmatrix}$$

## Result

(a)

$$X = \begin{bmatrix} -\frac{4}{3}t \\ -\frac{1}{2}s + \frac{1}{6}t \\ s \\ t \end{bmatrix}$$

(b) No solution (c)

$$X = \begin{bmatrix} \frac{2}{3} - \frac{4}{3}t \\ -\frac{1}{3} - \frac{1}{2}s + \frac{1}{6}t \\ s \\ t \end{bmatrix}$$

3. a

Given  $x_1 + x_2 + 2x_3 - x_4 = 3$ , here we are free to choose  $x_2, x_3$  and  $x_4$ . So take  $x_2 = s, x_3 = t$ , and  $x_4 = u$ , we get

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -s - 2t + u + 3 \\ s \\ t \\ u \end{bmatrix}$$

## Result

2 of 2

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -s - 2t + u + 3 \\ s \\ t \\ u \end{bmatrix}$$

4. a

Given

$$[A|I] = \left[ \begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ 2 & 6 & 0 & 1 \end{array} \right]$$

Now first operation is  $R_2 \rightarrow R_2 - 2R_1$

$$\left[ \begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ 2 & 6 & 0 & 1 \end{array} \right] \rightarrow \left[ \begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ 0 & -4 & -2 & 1 \end{array} \right]$$

which is given by

$$A_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

Now,

$$\left[ \begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ 0 & -4 & -2 & 1 \end{array} \right] \rightarrow \left[ \begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{4} \end{array} \right]$$

Here  $R_2 \rightarrow -\frac{R_2}{4}$  for operation is given by

$$A_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$$

The third operator is

$$\left[ \begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{4} \end{array} \right] \rightarrow \left[ \begin{array}{cc|cc} 1 & 0 & -\frac{3}{2} & \frac{5}{4} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{4} \end{array} \right]$$

in which  $R_1 \rightarrow R_1 - 5R_2$ , in which the operator is given by

$$A_3 = \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix}$$

Now,

$$\begin{aligned} A_3 A_2 A_1 &= \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{3}{2} & \frac{5}{4} \\ \frac{1}{2} & -\frac{1}{4} \end{bmatrix} \end{aligned}$$

which is same as we obtained by row reduction on augmented matrix.

## Result

Calculate

$$A_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

,

$$A_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$$

and

$$A_3 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

and show that

$$A_3 A_2 A_1 = \begin{bmatrix} -\frac{3}{2} & \frac{5}{4} \\ \frac{1}{2} & -\frac{1}{4} \end{bmatrix}$$

5. a

Given

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

To find  $A^{-1}$  we proceed by augmenting by  $I$ ,

$$\left[ \begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{R_1 \leftrightarrow R_2} \left[ \begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array} \right]$$

This shows that inverse of given matrix is itself. i.e.

$$A^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Given

$$B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$$

To find  $B^{-1}$  we proceed by augmenting by  $I$ ,

$$\begin{aligned} \left[ \begin{array}{cc|cc} 3 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array} \right] &\xrightarrow{R_1 \leftrightarrow R_2} \left[ \begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 3 & 5 & 1 & 0 \end{array} \right] \xrightarrow{R_2 \rightarrow R_2 - 3R_1} \left[ \begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & -3 \end{array} \right] \\ &\xrightarrow{R_2 \rightarrow -R_2} \left[ \begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 3 \end{array} \right] \xrightarrow{R_1 \rightarrow R_1 - 2R_2} \left[ \begin{array}{cc|cc} 1 & 0 & 2 & -5 \\ 0 & 1 & -1 & 3 \end{array} \right] \end{aligned}$$

This gives

$$B^{-1} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$$



Let

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Then

$$\begin{aligned} & \left[ \begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{R_2 \rightarrow R_2 - R_1} \left[ \begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array} \right] \\ & \xrightarrow{R_2 \rightarrow -R_2} \left[ \begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array} \right] \xrightarrow{R_1 \rightarrow R_1 - R_2} \left[ \begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \end{array} \right] \end{aligned}$$

Thus

$$C^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

Since

$$(CAB)(B^{-1}A^{-1}C^{-1}) = I$$

Thus, we get

$$\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix}$$

**Result**

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

,

$$\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$$

and

$$\begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix}$$

6. a

Given matrix

$$\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

Augmenting with identity matrix  $I_5$ , we get

$$\left[ \begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

Using row operation  $R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - R_1, R_4 \rightarrow R_4 - R_1$  and  $R_5 \rightarrow R_5 - R_1$ .

$$\left[ \begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 3 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 4 & 6 & 4 & 1 & -1 & 0 & 0 & 0 & 1 \end{array} \right]$$

Using row operation  $R_3 \rightarrow R_3 - 2R_2, R_4 \rightarrow R_4 - 3R_2$  and  $R_5 \rightarrow R_5 - 4R_2$ , we get

$$\left[ \begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & 0 & 6 & 4 & 1 & 3 & -4 & 0 & 0 & 1 \end{array} \right]$$

Using row operation  $R_4 \rightarrow R_4 - 3R_3$  and  $R_5 \rightarrow R_5 - 6R_3$

$$\left[ \begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 3 & -3 & 1 & 0 \\ 0 & 0 & 0 & 4 & 1 & -3 & 8 & -6 & 0 & 1 \end{array} \right]$$

Using row operation  $R_5 \rightarrow R_5 - 4R_4$ , we get

$$\left[ \begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 3 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -4 & 6 & -4 & 1 \end{array} \right]$$

Therefore the inverse of given pascal matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}$$

## Result

The inverse of given pascal triangle is

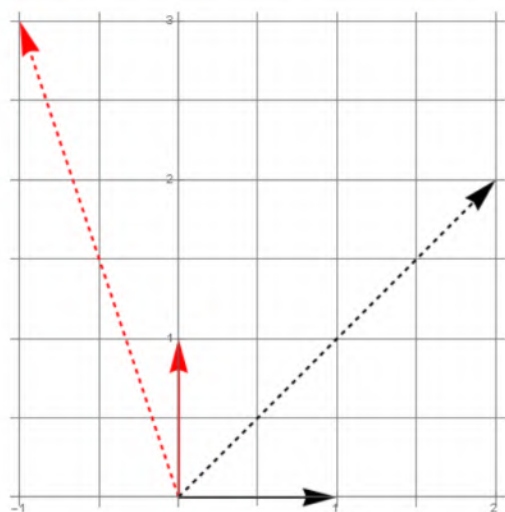
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}$$

7. a

When matrix

$$A = \begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}$$

acts by multiplication on unit vectors along axes, it transforms vectors as given in below picture. Any vector can be decomposed into components of unit vector along the axes. When  $A$  acts on the vector then it is the resultant of scaled components of the transformed unit vectors.



## Result

2 of 2

The sketch is plotted for unit vectors in the answer. Any general vector can be decomposed into unit vectors along which can be transformed by  $A$  as in image. The resultant vector is the transformed vector by  $A$ .

8. a

Given that  $AB$  is  $n \times n$  matrix which is invertible. Let  $C$  be it's multiplicative inverse i.e.  $CAB = ABC = I$  then  $(CA)B = I$  and  $A(BC) = I$ . We only need to show that  $B(CA) = I$  and  $(BC)A = I$ .

$$(CA)B = I \implies (CA)B(CA) = I(CA) \implies (CA)B(CA)B = I(CA)B$$

Now using associativity of matrix multiplication, we get

$$(CA)B(CAB) = I(CAB) \implies (CA)B = I$$

Hence  $CA$  is multiplicative inverse of  $B$ . Similarly, we can show  $BC$  is multiplicative inverse of  $A$ . Moreover, we can show this using determinant property of invertible matrix. Since  $\det(AB) = \det(A)\det(B)$  and  $AB$  invertible implies  $\det(AB) \neq 0$  which implies  $\det(A) \neq 0$  and  $\det(B) \neq 0$  which in turn implies both  $A$  and  $B$  are invertible.

## Result

2 of 2

Let  $C$  be inverse of  $AB$  then show  $CA$  is inverse of  $B$ .

9. a

(a) Given that  $AX = B$  has more than one solution where  $A$  and  $B$  have only real components. Let  $X_1$  and  $X_2$  be two solutions of  $AX = B$ . Then  $AX_1 = AX_2 = B$ . Let  $s, t$  be any number such that  $s + t = 1$  then

$$\begin{aligned} A(sX_1 + tX_2) &= AsX_1 + AtX_2 \\ &= sAX_1 + tAX_2 \\ &= sB + tB \\ &= (s + t)B \\ &= B \end{aligned}$$

This shows that  $sX_1 + tX_2$  is also solution of  $AX = B$ . Since there are infinite numbers  $s, t$  with  $s + t = 1$ , there are infinite solutions for the system.

(b) Suppose  $AX = B$  has solution is complex number. Since each component of  $X$  has real and imaginary part, let  $X'$  be real part and  $X''$  be the imaginary part so

$$X = X' + iX''$$

Now, operating on both sides by  $A$ , we get

$$AX = AX' + iAX'' = B \quad (1)$$

Since both  $X', X''$  and  $B$  have only real parts, comparing both sides in (1), we conclude that  $X'' = 0$  i.e.  $AX' = B$  hence the system also has real solution.

## Result

3 of 3

(a) Take any two solutions and show that it's convex combination is also solution. (b) Show that the complex component of solution must be zero.

10. a

Let  $A$  be a square matrix and system  $AX = B$  has unique solution for some particular column vector  $B_1$ . Suppose for some  $B$ , it has two distinct solutions, say  $X_1$  and  $X_2$  then

$$A(X_1 - X_2) = AX_1 - AX_2 = B - B = 0$$

Let  $X'$  be unique solution for  $AX = B_1$  then since  $X' + X_1 - X_2 \neq X'$ , we get

$$A(X' + X_1 - X_2) = AX' + A(X_1 - X_2) = B + 0 = B$$

which shows that  $AX = B_1$  does not have unique solution contradicting our assumption. Thus  $AX = B$  must have unique solution for every  $B$ .

## Result

2 of 2

If  $AX = B$  has unique solution then row reducing  $A$  must transform into identity matrix since any matrix with zero pivot implies infinitely many solution.

## Section 3

1. a

Given a square matrix  $B$ , we have  $(B^t)^t = B$ . Using property  $(AB)^t = B^t A^t$  gives

$$(BB^t)^t = (B^t)^t B^t = BB^t$$

which shows that  $BB^t$  is symmetric matrix. Now, using property  $(A + B)^t = A^t + B^t$ ,

$$(B + B^t)^t = B^t + (B^t)^t = B^t + B = B + B^t$$

which shows that  $B + B^t$  is symmetric. To show that  $(A^{-1})^t = (A^t)^{-1}$ ,

$$(A^t(A^{-1})^t)^t = ((A^{-1})^t)^t (A^t)^t = A^{-1}A = I$$

Similarly,

$$((A^{-1})^t A^t)^t = (A^t)^t ((A^{-1})^t)^t = AA^{-1} = I$$

Since  $I^t = I$ , this shows that  $(A^t)^{-1} = (A^{-1})^t$ .

## Result

Use property  $(AB)^t = B^t A^t$ ,  $(A + B)^t = A^t + B^t$

2. a

Given that  $A$  and  $B$  are two symmetric matrices (i.e.  $A^t = A$  and  $B^t = B$ ) and suppose if  $AB = BA$  then

$$(AB)^t = (BA)^t = A^t B^t = AB$$

which shows that  $AB$  is symmetric matrix. Now suppose that  $AB$  is symmetric matrix,

$$(AB)^t = AB \implies B^t A^t = AB \implies BA = AB$$

Thus  $AB$  being symmetric implies  $AB = BA$ .

## Result

2 of 2

$$(AB)^t = (BA)^t = A^t B^t = AB$$

and

$$(AB)^t = AB \implies B^t A^t = AB \implies BA = AB$$

### 3. a

Row operation and then column operation on  $A$  can be represented as  $(E_2(E_1 A))^t$  where  $E_1$  is row operation and  $E_2$  is column operation. Now if we do the column operation first then we get  $(E_2 A^t)^t$ . Applying row operation gives us  $E_1(E_2 A^t)^t$ . Simplifying  $(E_2(E_1 A))^t$  gives us

$$(E_2(E_1 A))^t = ((E_1 A)^t)(E_2)^t = E_1 A(E_2)^t$$

and simplifying  $E_1(E_2 A^t)^t$  gives us

$$E_1(E_2 A^t)^t = E_1 A(A^t)^t(E_2)^t = E_1 A(E_2)^t$$

So we get no change at all.

## Result

2 of 2

The result is same.

### 4. a

If both row and column operations are allowed in simplification of matrix and when pivot point is not zero, all elements on it's corresponding row and column can be made zero. Thus, we end up with matrix where no zero elements only occur on the diagonal of the matrix.

## Result

2 of 2

A non-zero pivot can make all elements on it's row and column zero.

## Section 4

### 1. a

(a)

$$\begin{aligned}
 \det \left( \begin{bmatrix} 1 & i \\ 2-i & 3 \end{bmatrix} \right) &= (1)(3) - (i)(2-i) \\
 &= 3 - 2i + i^2 \\
 &= 3 - 2i - 1 \\
 &= 2 - 2i
 \end{aligned}$$

(b)

$$\begin{aligned}
 \det \left( \begin{bmatrix} 1 & 1 \\ 1 & -i \end{bmatrix} \right) &= (1)(-i) - (1)(1) \\
 &= -i - 1
 \end{aligned}$$

(c)

$$\begin{aligned}
 \det \left( \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \right) &= 2 \det \left( \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \right) + 1 \det \left( \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \\
 &= 4 - 1 \\
 &= 3
 \end{aligned}$$

(d)

$$\begin{aligned}
 \det \left( \begin{bmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 8 & 6 & 3 & 0 \\ 0 & 9 & 7 & 4 \end{bmatrix} \right) &= 1 \det \left( \begin{bmatrix} 2 & 0 & 0 \\ 6 & 3 & 0 \\ 9 & 7 & 4 \end{bmatrix} \right) \\
 &= 2 \det \left( \begin{bmatrix} 3 & 0 \\ 7 & 4 \end{bmatrix} \right) \\
 &= 2(12) \\
 &= 24
 \end{aligned}$$

**Result**(a)  $2 - 2i$  (b)  $-i - 1$  (c) 3 (d) 24

2. a

Given

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 5 & -2 \end{bmatrix}$$

Then

$$\det(A) = 2 \cdot 4 - 1 \cdot 3 = 5$$

$$\det(B) = 1(-2) - 5 \cdot 1 = -7$$

Now,

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 5 & -2 \end{bmatrix} = \begin{bmatrix} 17 & -4 \\ 21 & -7 \end{bmatrix}$$

Now

$$\det(AB) = 17(-7) - 21(-4) = -35$$

This shows  $\det(A) \det(B) = 5(-7) = -35 = \det(AB)$ .

### Result

Show that  $\det(AB) = \det(A) \det(B) = -35$

### 3. a

Suppose

$$A_n = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & \dots & \\ & & & \dots & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$$

be  $n \times n$  matrix.

Now,  $\det(A_1) = \det([2]) = 2$

$$\det(A_2) = \det \left( \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \right) = 3$$

Similarly,

$$\det(A_3) = \det \left( \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \right) = 4$$



We conjecture that  $\det(A_n) = n + 1$ . Suppose this is true for some positive integer  $k$  i.e.  $\det(A_k) = k + 1$

$$\begin{aligned}
 \det(A_{k+1}) &= \det \left( \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & \cdots & \\ & & & \cdots & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}_{(k+1) \times (k+1)} \right) \\
 &= 2 \det \left( \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & \cdots & \\ & & & \cdots & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}_{k \times k} \right) \\
 &\quad + \det \left( \begin{bmatrix} -1 & -1 & & & \\ 0 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & \cdots & \\ & & & \cdots & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}_{k \times k} \right) \\
 &= 2(k+1) - \det \left( \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & \cdots & \\ & & & \cdots & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}_{(k-1) \times (k-1)} \right) \\
 &= 2(k+1) - k \\
 &= k+2
 \end{aligned}$$

## Result

2 of 2

If  $\det(A_k) = k + 1$  and  $\det(A_{k-1}) = k$  then show  $\det(A_{k+1}) = k + 2$ . Then by induction,  $\det(A_n) = n + 1$  for all positive integers.

4. a

Now, let  $I_n$  be  $n \times n$  identity matrix. Then

$$\det(-I_n) = \det \left( \begin{bmatrix} -1 & & & & \\ & -1 & & & \\ & & -1 & & \\ & & & \cdots & \\ & & & & -1 & \\ & & & & & -1 \end{bmatrix}_{n \times n} \right) = (-1)^n$$

Thus,

$$\det(-A) = \det((-I_n)A) = \det(-I_n) \det(A) = (-1)^n \det(A)$$

## Result

$$\det(-A) = (-1)^n \det(A)$$

5. a

We consider two cases, when  $\det A = 0$  and when  $\det A \neq 0$ . When  $\det A = 0$ ,  $A$  is not invertible therefore  $(A^t)^{-1} = (A^{-1})^t$  implies  $A^t$  is also not invertible so  $\det(A^t) = 0$  which gives  $\det A = \det(A^t)$ .

Now if  $\det A \neq 0$  implies  $A$  is invertible hence there exist elementary matrices of type (i) and (iii) given by (1.2.5) such that

$$A = E_1 E_2 \dots E_n I$$

Since  $E_k$  are of type (i) or (iii),  $\det E_k = \det((E_k)^t)$  since for type (i), we have  $\det(E_k) = 1$  and  $(E_k)^t = E_k$  for type (iii). Thus, we get

$$A^t = I E_n^t \dots (E_2)^t (E_1)^t$$

which gives

$$\begin{aligned} \det(A^t) &= \det(E_n^t) \dots \det(E_2^t) \det(E_1^t) \\ &= \det(E_n) \dots \det(E_2) \det(E_1) \\ &= \det(E_1) \det(E_2) \dots \det(E_n) \\ &= \det(E_1 E_2 \dots E_n I) \\ &= \det(A) \end{aligned}$$

## Result

2 of 2

If  $\det A \neq 0$ , show that it can be written as product of elementary matrices of type (i) and (iii) for which  $\det E = \det(E^t)$ .

## 6. a

First of all square matrix  $A$ , we will show that

$$\det \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} = \det(A) \quad (1)$$

and

$$\det \begin{bmatrix} I & B \\ 0 & D \end{bmatrix} = \det(D) \quad (2)$$

Let  $A$  be  $m \times m$  matrix. To show (1), we will proceed by induction. Suppose  $B_1$  is  $1 \times 1$  matrix then

$$\det \begin{bmatrix} A & b_{11} \\ 0 & 1 \end{bmatrix} = (-1)^2 \det(A) = \det(A)$$

Suppose the result is true for  $(n-1) \times (n-1)$  matrix with  $B_{n-1}$  being any  $(n-1) \times (n-1)$  matrix, i.e.

$$\det \begin{bmatrix} A & B \\ 0 & I_{n-1} \end{bmatrix} = \det(A)$$

then for any  $n \times n$  matrix  $B_n$ ,

$$\begin{aligned}
\det \begin{bmatrix} A & B_{n+1} \\ 0 & I_{n+1} \end{bmatrix} &= \left[ \begin{array}{c|ccc} A & b_{11} & \dots & b_{1n} \\ \vdots & \vdots & \dots & \vdots \\ b_{m1} & \dots & \dots & b_{mn} \\ \hline & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 1 \end{array} \right] \\
&= \left[ \begin{array}{c|ccc} A & b_{11} & \dots & b_{1(n-1)} \\ \vdots & \vdots & \dots & \vdots \\ b_{m1} & \dots & \dots & b_{m(n-1)} \\ \hline & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 1 \end{array} \right] \\
&= \det \begin{bmatrix} A & B_{n-1} \\ 0 & I_{n-1} \end{bmatrix} \\
&= \det A
\end{aligned}$$

Thus by induction (1) is true for all  $I_n$ . We similarly proceed to show (2) is true as well.

$$\begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \times \begin{bmatrix} I & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} A & AB + BD \\ 0 & D \end{bmatrix} = \begin{bmatrix} A & B' \\ 0 & D \end{bmatrix}$$

where  $B' = AB + BD$ .

$$\begin{aligned}
\det \begin{bmatrix} A & B' \\ 0 & D \end{bmatrix} &= \det \left( \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \times \begin{bmatrix} I & B \\ 0 & D \end{bmatrix} \right) \\
&= \det \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \times \det \begin{bmatrix} I & B \\ 0 & D \end{bmatrix} \\
&= \det(A) \det(D)
\end{aligned}$$

This can also be shown using elementary matrices. If  $\det A = 0$  or  $\det D = 0$  then simplification of the block matrix gives zero in one of pivot for which determinant is zero. If it is not the case, let  $A = E_1 \dots E_n$ ,  $B = F_1 F_2 \dots F_m$  then take

$$E'_k = \begin{bmatrix} E_k & 0 \\ 0 & I \end{bmatrix}, \quad F'_k = \begin{bmatrix} I & 0 \\ 0 & F_k \end{bmatrix}$$

then  $\det E_k = \det E'_k$  and  $\det F_k = \det F'_k$ . Now  $E'_1, \dots, E'_n F'_1 \dots, F'_m$  are elementary matrices for the block matrix which simplifies it i.e.

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = E'_1 E'_2 \dots E'_n F'_1 F'_2 \dots F'_m$$

which gives

$$\begin{aligned}
 \det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} &= \det(E'_1 E'_2 \dots E'_n F'_1 F'_2 \dots F'_m) \\
 &= \det(E'_1) \dots \det(E'_n) \det(F'_1) \dots \det(F'_m) \\
 &= \det(E_1) \dots \det(E_n) \det(F_1) \dots \det(F_m) \\
 &= \det(E_1 \dots E_n) \det(F_1 \dots F_m) \\
 &= \det A \det B
 \end{aligned}$$

## Result

3 of 3

The result can be proved using induction as well as using elementary matrices. For elementary matrices of  $A$  and  $D$ , redefine new elementary matrices which simplifies block matrices. Taking determinant gives the required result.

## Section 5

1. a

Given  $(1\ 2)(1\ 3)(1\ 4)(1\ 5)$ .

$$\begin{aligned}
 (1\ 2)(1\ 3)(1\ 4)(1\ 5) &\stackrel{(1)}{=} (1\ 2)(1\ 3)(5\ 4\ 1) \\
 &\stackrel{(2)}{=} (1\ 2)(1\ 3)(1\ 5\ 4) \\
 &\stackrel{(3)}{=} (1\ 2)(4\ 3\ 1\ 5) \\
 &\stackrel{(4)}{=} (1\ 2)(1\ 5\ 4\ 3) \\
 &\stackrel{(5)}{=} (3\ 2\ 1\ 5\ 4) \\
 &\stackrel{(6)}{=} (1\ 5\ 4\ 3\ 2)
 \end{aligned}$$

Explanations,

1.  $(1\ 4)(1\ 5) = (5\ 4\ 1)$
2.  $(5\ 4\ 1) = (1\ 5\ 4)$
3.  $(1\ 3)(1\ 5\ 4) = (4\ 3\ 1\ 5)$
4.  $(4\ 3\ 1\ 5) = (1\ 5\ 4\ 3)$
5.  $(1\ 2)(1\ 5\ 4\ 3) = (3\ 2\ 1\ 5\ 4)$
6.  $(3\ 2\ 1\ 5\ 4) = (1\ 5\ 4\ 3\ 2)$

Given  $(1\ 2\ 3)(2\ 3\ 4)(3\ 4\ 5)$ ,

$$\begin{aligned}
 (1\ 2\ 3)(2\ 3\ 4)(3\ 4\ 5) &\stackrel{(1)}{=} (1\ 2\ 3)(3\ 2)(5\ 4) \\
 &= (1\ 2\ 3)(2\ 3)(4\ 5) \\
 &\stackrel{(2)}{=} (2\ 1)(4\ 5) \\
 &= (1\ 2)(4\ 5)
 \end{aligned}$$

Explanations,

1.  $(2\ 3\ 4)(3\ 4\ 5) = (2\ 3)(4\ 5)$ .
2.  $(1\ 2\ 3)(2\ 3) = (1\ 2)$

Given  $(1\ 2\ 3\ 4)(2\ 3\ 4\ 5)$ ,

$$\begin{aligned}
 (1\ 2\ 3\ 4)(2\ 3\ 4\ 5) &\stackrel{(1)}{=} (1\ 4)(1\ 3)(1\ 2)(2\ 3\ 4\ 5) \\
 &\stackrel{(2)}{=} (1\ 4)(1\ 3)(1\ 2\ 3\ 4\ 5) \\
 &\stackrel{(3)}{=} (1\ 4)(2\ 1)(3\ 4\ 5) \\
 &= (1\ 4)(1\ 2)(3\ 4\ 5) \\
 &\stackrel{(4)}{=} (1\ 2\ 4)(3\ 4\ 5) \\
 &\stackrel{(5)}{=} (3\ 1\ 2\ 4\ 5) \\
 &= (1\ 2\ 4\ 5\ 3)
 \end{aligned}$$

Explanations,

1.  $(1\ 2\ 3\ 4) = (1\ 4)(1\ 3)(1\ 2)$ .
2.  $(1\ 2)(2\ 3\ 4\ 5) = (1\ 2\ 3\ 4\ 5)$ .
3.  $(1\ 3)(1\ 2\ 3\ 4\ 5) = (2\ 1)(3\ 4\ 5)$ .
4.  $(1\ 4)(1\ 2) = (1\ 2\ 4)$
5.  $(1\ 2\ 4)(3\ 4\ 5) = (3\ 1\ 2\ 4\ 5)$

Given  $(1\ 2)(2\ 3)(3\ 4)(4\ 5)(5\ 1)$ ,

$$\begin{aligned}
 (1\ 2)(2\ 3)(3\ 4)(4\ 5)(5\ 1) &\stackrel{(1)}{=} (1\ 2)(2\ 3)(3\ 4)(1\ 4\ 5) \\
 &\stackrel{(2)}{=} (1\ 2)(2\ 3)(3\ 5\ 1\ 4) \\
 &= (1\ 2)(2\ 3)(1\ 4\ 3\ 5) \\
 &\stackrel{(3)}{=} (1\ 2)(4\ 2\ 3\ 5\ 1) \\
 &= (1\ 2)(1\ 4\ 2\ 3\ 5) \\
 &\stackrel{(4)}{=} (5\ 2\ 3)(1\ 4) \\
 &= (2\ 3\ 5)(1\ 4)
 \end{aligned}$$

Explanations,

1.  $(4\ 5)(5\ 1) = (1\ 4\ 5)$ .
2.  $(3\ 4)(1\ 4\ 5) = (3\ 5\ 1\ 4)$ .
3.  $(2\ 3)(1\ 4\ 3\ 5) = (4\ 2\ 3\ 5\ 1)$ .
4.  $(1\ 2)(1\ 4\ 2\ 3\ 5) = (5\ 2\ 3)(1\ 4)$ .

## Result

$$(1\ 5\ 4\ 3\ 2), (1\ 2)(4\ 5), (1\ 2\ 4\ 5\ 3), (2\ 3\ 5)(1\ 4)$$

2. a

(a) The associated matrix with the permutation  $(1\ 3\ 4\ 2)$  is

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(for given row, the column given by it's successor in  $(1\ 3\ 4\ 2)$  is 1)

(b) The permutation  $(1\ 3\ 4\ 2)$  can also be written as

$$(1\ 3\ 4\ 2) = (1\ 2)(1\ 4)(1\ 3)$$

The associated permutation matrix is

$$(1\ 2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$(1\ 4) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$(1\ 3) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus, we get

$$(1\ 2)(1\ 4)(1\ 3) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

which is same as the result we got in (a)

(c) The sign of  $p$  is  $-1$  since it's odd permutation because it consists of 3 (which is odd) permutations. The determinant of associated matrix is also  $-1$ .

## Result

(a)

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$(b) (1\ 3\ 4\ 2) = (1\ 2)(1\ 4)(1\ 3) \quad (c) -1$$

3. a

Let  $E = (a_{ij})$  be an elementary permutation matrix which permutes two elements of a row  $k \leftrightarrow l$ . Then

$$a_{ij} = \begin{cases} 1 & \text{if } i = l, j = k, \\ & \text{or } i = k, j = l \\ & \text{or } i = j \neq (l = k) \\ 0 & \text{otherwise} \end{cases}$$

Then  $EE^t = (b_{ij})$  where  $b_{ij} = \sum_m a_{im}a_{mj}$ . If  $i = l, m = k$  then  $a_{im} = 1$ . If  $a_{mj} = 1$  then since  $m = k$ , we must have  $i = l$  i.e.  $a_{ij} = 1$  if  $i = j = l$  or  $i = j = k$ . Similarly  $i = k, m = l$ , we get  $i = j$  if  $b_{ij} = 1$ . If that is not the case then  $i = m = j$  which gives  $a_{ij} = 1$  if  $i = j$ . Thus,  $b_{ij} = 1$  if  $i = j$ . Hence

$$EE^t = I$$

Since any permutation can be written as product of elementary transposition matrices,

$$P = E_1 E_2 \dots E_n$$

, we get

$$\begin{aligned} PP^t &= (E_1 E_2 \dots E_n)(E_1 E_2 \dots E_n)^t \\ &= E_1 E_2 \dots E_n E_n^t \dots E_2^t E_1^t \\ &= I \end{aligned}$$

and

$$\begin{aligned} P^t P &= (E_1 E_2 \dots E_n)^t (E_1 E_2 \dots E_n) \\ &= E_n^t \dots E_2^t E_1^t E_1 E_2 \dots E_n \\ &= I \end{aligned}$$

Hence, the transpose of elementary matrix is its inverse.

## Result

2 of 2

Show that transpose of elementary transposition matrix is its inverse. Then use the fact that each permutation can be written as product of elementary transposition matrices to show the required result.

4. a

Given  $p(i) = n - i + 1$  is permutation of  $n$  indices. We have

$$p(n - i + 1) = n - (n - i + 1) + 1 = i$$

Thus if  $P = (a_{ij})$  is the associated permutation matrix, we get  $a_{ij} = 1$  if  $j = n - i + 1$  otherwise 0.

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} \quad (1)$$

If  $n$  is even then

$$p\left(\frac{n}{2}\right) = n - \frac{n}{2} + 1 = \frac{n}{2} + 1$$

$$p\left(\frac{n}{2} - 1\right) = n - \frac{n}{2} - 1 + 1 = \frac{n}{2}$$

If  $n$  is odd then

$$p\left(\frac{n+1}{2}\right) = n - \frac{n+1}{2} + 1 = \frac{n+1}{2}$$

$$p\left(\frac{n-1}{2}\right) = n - \frac{n-1}{2} + 1 = \frac{n+3}{2}$$

Thus if  $n$  is even then cycle decomposition is given by

$$p = (1 \ n)(2 \ (n-1)) \dots \left(\left(\frac{n}{2}\right) \ \left(\frac{n}{2} + 1\right)\right) \quad (2)$$

If  $n$  is odd then the cycle decomposition is given by

$$p = (1 \ n)(2 \ (n-1)) \dots \left(\left(\frac{n-1}{2}\right) \ \left(\frac{n+3}{2}\right)\right) \quad (3)$$

Now for even  $n$ , if  $n/2$  is odd then there are odd number of transpositions sign is  $-1$ . If  $n/2$  is even then there are even number of transpositions so sign is  $+1$ .

Similarly if  $n$  is odd then if  $(n-1)/2$  is odd then sign is  $-1$  and if  $(n-1)/2$  is even then sign is  $+1$ .

## Result

2 of 2

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

The sign of permutation is  $+1$  if  $n \equiv 0, 1 \pmod{4}$  and  $-1$  if  $n \equiv 2, 3 \pmod{4}$ .

5. a



Given  $p = (3\ 4\ 1)(2\ 5)$  and  $q = (1\ 4\ 5\ 2)$  and  $pq = (2\ 3\ 4)$  and  $qp = (1\ 3\ 5)$ . If the result is not same, it is no strange as the permutation operation is not commutative. Both results turn out to be 3-cycles. However, this is no accident since

$$(1\ 2\ 3\ 5\ 4)(2\ 3\ 4) = (1\ 3\ 5)(1\ 2\ 3\ 5\ 4)$$

i.e.

$$(1\ 2\ 3\ 5\ 4)pq = qp(1\ 2\ 3\ 5\ 4) \implies qp = g(qp)g^{-1}$$

where  $g = (1\ 2\ 3\ 5\ 4)$ . Now, if  $(pq)(i) = j$  then

$$g(pq)g^{-1}(g(i)) = g((pq)(i)) = g(j)$$

Now if  $pq$  is a cycle of length 3 then  $(qp)(j) = g((pq)(i))$ . This shows if  $pq$  is cycle of length 3,  $qp$  is also cycle of length 3.

## Result

2 of 2

$pq$  and  $qp$  happen to be conjugates to each other.

## Section 6

1. a

(a)

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = (-1)^{1+2}3 \cdot 2 + (-1)^{2+2}4 \cdot 1 = -2$$

$$\begin{aligned} \det \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix} &= (-1)^{3+1}(0) \det \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} + (-1)^{3+2}(2) \det \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \\ &\quad + (-1)^{3+3}(1) \det \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \\ &= 0 - 2(1 \cdot 2 - 2 \cdot 2) + (4 \cdot 1 - 1 \cdot 2) \\ &= 4 + 2 \\ &= 6 \end{aligned}$$

$$\begin{aligned} \det \begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} &= (-1)^{3+1}(1) \det \begin{bmatrix} b & c \\ 0 & 1 \end{bmatrix} + (-1)^{3+2}(1) \det \begin{bmatrix} a & c \\ 1 & 1 \end{bmatrix} \\ &\quad + (-1)^{3+3}(1) \det \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \\ &= (b \cdot 1 - 0 \cdot c) - (a \cdot 1 - c \cdot 1) + (a \cdot 0 - 1 \cdot b) \\ &= b - a + c - b \\ &= c - a \end{aligned}$$

(b) Using complete expansion, the determinant of matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is given by

$$\begin{aligned} \det A &= \sum_{\text{perm } p} (\text{sign } p) a_{1,p1} \cdots a_{n,pn} \\ &= a_{11}a_{22} - a_{21}a_{12} \end{aligned}$$

Now,

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1 \cdot 4 - 3 \cdot 2 = -2$$

Using complete expansion, the determinant of matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

is given by

$$\begin{aligned} \det A &= \sum_{\text{perm } p} (\text{sign } p) a_{1,p1} \cdots a_{n,pn} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{33} \end{aligned}$$

Now,

Using complete expansion, the determinant of matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is given by

$$\begin{aligned} \det A &= \sum_{\text{perm } p} (\text{sign } p) a_{1,p1} \cdots a_{n,pn} \\ &= a_{11}a_{22} - a_{21}a_{12} \end{aligned}$$

Now,

$$\begin{aligned} \det \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix} &= 1 \cdot 4 \cdot 1 + 1 \cdot 2 \cdot 0 + 2 \cdot 2 \cdot 2 \\ &\quad - 1 \cdot 2 \cdot 2 - 1 \cdot 2 \cdot 1 - 2 \cdot 4 \cdot 0 \\ &= 4 + 8 - 4 - 2 \\ &= 6 \end{aligned}$$

Similarly,

$$\begin{aligned} \det \begin{bmatrix} 4 & -1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix} &= 4 \cdot 1 \cdot 1 + (-1) \cdot (-2) \cdot 1 + 1 \cdot 1 \cdot (-1) \\ &\quad - 4 \cdot (-2) \cdot (-1) - (-1) \cdot 1 \cdot 1 - 1 \cdot 1 \cdot 1 \\ &= 4 + 2 - 1 - 8 + 1 - 1 \\ &= -3 \end{aligned}$$

Similarly,

$$\begin{aligned} \det \begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} &= a \cdot 0 \cdot 1 + b \cdot 1 \cdot 1 + c \cdot 1 \cdot 1 \\ &\quad - a \cdot 1 \cdot 1 - b \cdot 1 \cdot 1 - c \cdot 0 \cdot 1 \\ &= b + c - a - b \\ &= c - a \end{aligned}$$

(c) Let  $A$  be  $n \times n$  matrix and  $\text{cof}(A)$  be its cofactor matrix. The  $\text{cof}(A)$  is  $n \times n$  matrix whose  $i, j$  entry is given by

$$\text{cof}(A)_{ij} = (-1)^{i+j} \det A_{ji}$$

Now, given

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Now,

$$\text{cof}(A)_{11} = (-1)^{1+1} \det A_{11} = 4$$

$$\text{cof}(A)_{22} = (-1)^{2+2} \det A_{22} = 1$$

$$\text{cof}(A)_{12} = (-1)^{1+2} \det A_{21} = -3$$

$$\text{cof}(A)_{21} = (-1)^{2+1} \det A_{12} = -2$$

Thus,

$$\text{cof}(A) = \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}$$

Similarly, we compute

$$\text{cof} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 4 \\ 3 & 1 & -2 \\ -6 & 2 & 2 \end{bmatrix}$$

$$\text{cof} \begin{bmatrix} 4 & -1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -3 & -2 \\ 0 & 3 & 3 \\ 1 & 9 & 5 \end{bmatrix}$$

$$\text{cof} \begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ c-b & a-c & b-a \\ b & c-a & -b \end{bmatrix}$$

## Result

6 c

The determinants of given matrices are computed as  $-2$ ,  $6$ ,  $-3$  and  $c - a$  using both row expansion and complete expansion. The cofactor matrices have been similarly computed in answer.

## 2. a

Given that  $A$  is invertible matrix. If  $A^{-1}$  as all integer entries  $a_{ij}$ , then since

$$\det A = \sum_{\text{permp}} (\text{sign } p) a_{1,p1} \cdots a_{n,pn}$$

We must have  $\det A$  and  $\det(A^{-1})$  as integers. Thus,

$$\det(AA^{-1}) = \det(A) \det(A^{-1}) = \det I = 1$$

Thus, if  $\det A = n$ , then  $\det(A^{-1}) = \frac{1}{n}$  which must also be an integer. For only values  $n = \pm 1$ , we get integer hence

$$\det A = \pm 1$$

Conversely, if  $\det A = \pm 1$ , then by Theorem 1.6.9  $A^{-1} = \text{cof}(A) / \det(A)$ . Since  $i, j$  entry of cofactor of  $A$  is given by

$$\text{cof}(A)_{ij} = (-1)^{i+j} \det A_{ji}$$

and also  $A$  contains only integer entries,  $\text{cof}(A)$  is also integer matrix. Thus,

$$A^{-1} = \frac{\text{cof } A}{\det A} = \pm \text{cof } A$$

which is also integer matrix.

## Result

3 of 3

Use the fact that  $\det(AA^{-1}) = \det(A) \det(A^{-1}) = 1$  to show that  $\det A = \pm 1$ . Conversely, use  $A^{-1} = \text{cof}(A) / \det(A)$  to show  $A^{-1}$  contains only integer entries.

## Miscellaneous Problems

1. a

Given  $2n \times 2n$  block matrix

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where each block is  $n \times n$  matrix. Suppose  $A$  is invertible and  $AC = CA$  then

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

Now, we know that

$$\det \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} = 1$$

so we get

$$\begin{aligned} \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \det \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix} \\ &= \det(A) \det(D - CA^{-1}B) \\ &= \det(AD - ACA^{-1}B) \\ &= \det(AD - CAA^{-1}B) \\ &= \det(AD - CB) \end{aligned}$$

The result does not hold if  $AC \neq CA$ . For example take

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$$

and  $B = D = I$ . This gives

$$\det M = \det \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & 3 & 4 \\ 0 & 1 & 1 & 2 \end{bmatrix} = -28$$

while  $\det(AD - CB) = \det(A - C) = 0$ .

---

### Result

Use the property

$$\det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \det(A) \det(D)$$

to show the required result.

2. a

Suppose that  $A$  has a left inverse. Denote it by  $B$ . By definition,  $B$  must be an  $n \times m$  matrix such that

$$BA = I_n$$

Let  $\tilde{A}$  be a matrix which we get when we add  $(n - m)$  rows of zeros to  $A$  at the bottom. Let  $\tilde{B}$  be a matrix which we get when we add  $(n - m)$  columns of zeros to  $B$  to the right. Notice that  $\tilde{A}$  and  $\tilde{B}$  are  $n \times n$  matrices, for which

$$\tilde{B}\tilde{A} = I_n$$

holds. However, we know that now the following also holds:

$$\tilde{A}\tilde{B} = I_n$$

(if a square matrix has a one-sided inverse, then the other inverse also exists and they are equal). So,  $\tilde{A}$  is invertible. However, this is impossible by Lemma 1.1.18.

## Result

2 of 2

Assume that the left inverse exists and arrive to a contradiction by a hint provided in the text of the exercise. Lemma 1.1.18. will prove useful.

### 3. a

Let

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{bmatrix}$$

Then

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{n1} + b_{n1} & \dots & a_{nn} + b_{nn} \end{bmatrix}$$

Therefore,

$$\begin{aligned} \text{trace}(A + B) &= a_{11} + b_{11} + a_{22} + b_{22} + \dots + a_{nn} + b_{nn} \\ &= a_{11} + a_{22} + \dots + a_{nn} + b_{11} + b_{22} + \dots + b_{nn} \\ &= \text{trace}(A) + \text{trace}(B) \end{aligned}$$

which proves the first statement.

Now,

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1n}b_{n1} & \dots & a_{11}b_{1n} + a_{12}b_{2n} + \dots + a_{1n}b_{nn} \\ \vdots & \ddots & \vdots \\ a_{n1}b_{11} + a_{n2}b_{21} + \dots + a_{nn}b_{n1} & \dots & a_{n1}b_{1n} + a_{n2}b_{2n} + \dots + a_{nn}b_{nn} \end{bmatrix}$$

Now notice that

$$\text{trace}(AB) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n a_{ij} b_{kl}$$

(written without these tedious sums, the trace of  $AB$  is equal to the sum of all combinations of products of entries of  $A$  and  $B$ ).

Similarly,

$$BA = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} + \dots + b_{1n}a_{n1} & \dots & b_{11}a_{1n} + b_{12}a_{2n} + \dots + b_{1n}a_{nn} \\ \vdots & \ddots & \vdots \\ b_{n1}a_{11} + b_{n2}a_{21} + \dots + b_{nn}a_{n1} & \dots & b_{n1}a_{1n} + b_{n2}a_{2n} + \dots + b_{nn}a_{nn} \end{bmatrix}$$

Now notice that

$$\text{trace}(BA) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n b_{kl} a_{ij} = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n a_{ij} b_{kl}$$

Finally,

$$\text{trace}(AB) = \text{trace}(BA)$$

Thus, the second statement is proven.

To prove the last statement, we use the previous statement:

$$\text{trace}(BAB^{-1}) = \text{trace}(B(AB^{-1})) = \text{trace}(AB^{-1}B) = \text{trace}(A)$$

With this, all statements are proven.

## Result

3 of 3

The first two statements are proved directly. The last statement can be easily proved using the second statement.

4. a

Here we will use a little trick. First assume that such  $A, B$  exist. Then we must have

$$\text{trace}(AB - BA) = \text{trace}(I)$$

On the left side, we have

$$\text{trace}(AB - BA) = \text{trace}(AB) - \text{trace}(BA) = 0,$$

by the previous exercise.

On the other hand,

$$\text{trace}(I) = \text{trace} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = n$$

Since  $0 \neq n$ , the equality  $\text{trace}(AB - BA) = \text{trace}(I)$  cannot hold, which means that the equality  $AB - BA = I$  too cannot hold.

## Result

Hint: take trace of both sides.

5. a

The given matrix can be written as

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

The expression is as short as possible. If it were possible to write it up as product of two elementary matrices,  $E_1 E_2 = M$  then

$$E_2^{-1} E_1^{-1} M = I$$

Since inverse of elementary matrix is elementary, this shows that it takes two elementary operation to reduce given matrix into reduced row echelon form. However it takes three elementary row operations to reduce it. Hence it cannot be written as product of two elementary matrices.

## Result

2 of 2

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

. To show it is shortest possible, reduce given matrix using inverse of elementary matrices.

6. a



Let  $A$  be any invertible  $2 \times 2$  matrix. Suppose

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

Take elementary matrix  $E_1$  such that

$$E_1 A = \begin{bmatrix} 1 & \\ & \end{bmatrix}$$

and  $E_2$  such that

$$E_2 E_1 A = \begin{bmatrix} 1 & \\ 0 & b_4 \end{bmatrix}$$

. Since  $A$  is invertible and  $E_1, E_2$  are elementary matrices which perform elementary row operations,  $b_4 \neq 0$ .

Now take  $E_3$  such that

$$E_3 E_2 E_1 A = \begin{bmatrix} 1 & \\ 0 & 1 \end{bmatrix}$$

. Now finally, we can take  $E_4$  such that  $E_4 E_3 E_2 E_1 A = I$ . Now, it gives

$$E_4 E_3 E_2 E_1 A = I \implies I = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1}$$

Since inverse of elementary matrix is elementary, every  $2 \times 2$  invertible matrix can be written as product of at most four elementary matrices.

## Result

$$n = 4$$

7. a

(a) Given

$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}$$

Multiplying by elementary matrix of type (i) as given by 1.4.13 does not change the determinant of any matrix. So

$$\begin{aligned} \det \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix} &= \det \begin{bmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b^2-a^2 & c^2-a^2 \end{bmatrix} \\ &= \det \begin{bmatrix} b-a & c-a \\ b^2-a^2 & c^2-a^2 \end{bmatrix} \\ &= (b-a)(c^2-a^2) - (c-a)(b^2-a^2) \\ &= (b-a)(c-a)(c-a-(b-a)) \\ &= (b-a)(c-a)(c-b) \\ &= (a-b)(c-b)(c-a) \end{aligned}$$



(b) We prove the formula for determinant using mathematical induction. Suppose let

$$\det \begin{bmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_n \\ \vdots & & \vdots \\ a_1^{n-1} & \dots & a_n^{n-1} \end{bmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

Now, doing  $R_{n+1} = R_{n+1} - a_1 R_n$ , we get

$$\begin{aligned} & \det \begin{bmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_{n+1} \\ \vdots & & \vdots \\ (a_1)^n & \dots & (a_{n+1})^n \end{bmatrix} \\ = & \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_{n+1} \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \dots & (a_{n+1})^{n-1} \\ 0 & (a_2)^n - a_1(a_2)^{n-1} & \dots & (a_{n+1})^n - a_1(a_{n+1})^{n-1} \end{bmatrix} \end{aligned}$$

Similarly, elementary row operation  $R_k \rightarrow R_k - a_1 R_{k-1}$  with  $k = n, n-1, \dots, 3, 2$  gives

$$= \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & a_2 - a_1 & \dots & a_{n+1} - a_1 \\ \vdots & \vdots & & \vdots \\ 0 & a_2^{n-1} - a_1(a_2)^{n-2} & \dots & (a_{n+1})^{n-1} - a_1(a_{n+1})^{n-2} \\ 0 & (a_2)^n - a_1(a_2)^{n-1} & \dots & (a_{n+1})^n - a_1(a_{n+1})^{n-1} \end{bmatrix}$$

So, we get

$$= \det \begin{bmatrix} a_2 - a_1 & \dots & a_{n+1} - a_1 \\ \vdots & & \vdots \\ a_2^{n-1} - a_1(a_2)^{n-2} & \dots & (a_{n+1})^{n-1} - a_1(a_{n+1})^{n-2} \\ (a_2)^n - a_1(a_2)^{n-1} & \dots & (a_{n+1})^n - a_1(a_{n+1})^{n-1} \end{bmatrix}$$

Upon factoring  $a_k - a_1$  from each row, we get

$$= (a_2 - a_1) \cdots (a_{n+1} - a_1) \det \begin{bmatrix} 1 & \cdots & 1 \\ a_2 & \cdots & a_{n+1} \\ \vdots & & \vdots \\ (a_2)^{n-1} & \cdots & (a_{n+1})^{n-1} \end{bmatrix}$$

Now, using induction hypothesis, the determinant is

$$\det \begin{bmatrix} 1 & \cdots & 1 \\ a_2 & \cdots & a_{n+1} \\ \vdots & & \vdots \\ (a_2)^{n-1} & \cdots & (a_{n+1})^{n-1} \end{bmatrix} = \prod_{2 \leq i < j \leq n+1} (a_j - a_i)$$

which gives the determinant as

$$\begin{aligned} \det \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_{n+1} \\ \vdots & & \vdots \\ (a_1)^n & \cdots & (a_{n+1})^n \end{bmatrix} &= (a_2 - a_1) \cdots (a_{n+1} - a_1) \prod_{2 \leq i < j \leq n+1} (a_j - a_i) \\ &= \prod_{1 \leq i < j \leq n+1} (a_j - a_i) \end{aligned}$$

(c) Consider the polynomial given by

$$P_k(t) = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ t & t_0 & \cdots & t_{k-1} & t_{k+1} & \cdots & t_n \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ t^n & t_0^n & \cdots & (t_{k-1})^n & (t_{k+1})^n & \cdots & (t_n)^n \end{bmatrix}$$

Now  $P_k(t_j) = 0$  if  $j \neq k$ . If  $(t_k, y_k)$  be any point then choose  $c_k$  such that  $c_k = y_k / P_k(t_k)$ . The  $n$  degree polynomial given by

$$P(t) = \sum_{k=0}^n c_k P_k(t)$$

is the polynomial that passes through any given  $(t_k, y_k)$ .

## Result

5 of 5

(a) Expanding the determinant gives the required result. (b) Use induction to show that determinant is given by  $\prod_{1 \leq i < j \leq n} (a_j - a_i)$ . (c) Such polynomial can be constructed by replacing each column by  $1, t, \dots, t^n$  in Vandermonde matrix. The sum of determinant of such matrix with appropriate coefficient gives the required polynomial.

8. a

(a)

This is provided in the text of the exercise.

(b)

This actually means that  $L$  must also be a right inverse! So, there is actually no problem here, and  $L$  is a two-sided inverse. Thus,  $A$  is invertible in the full sense.

## Result

2 of 2

This actually means that  $L$  must also be a right inverse! So, there is actually no problem here, and  $L$  is a two-sided inverse. Thus,  $A$  is invertible in the full sense.

9. a

Given  $A$  is  $2 \times 2$  matrix and  $A_1, A_2$  be the columns of  $A$ . Now if  $P$  is a parallelogram whose vertices are  $0, A_1, A_2, A_1 + A_2$  then the area of  $P$  remains unchanged under elementary row operation in which rows are swapped and multiple of rows are added to other. This represents the multiplication to  $A$  by elementary matrix of type first and second given by 1.4.13. The row operation in which a row is scaled by  $c$  also scales the area of the parallelogram.

The parallelogram with vertices  $(0, 0), (1, 0), (0, 1), (1, 1)$  have column  $(1, 0)$  and  $(0, 1)$  forms the matrix  $I$ . The area is also 1 which is same as  $\det I$ . Now for any invertible given matrix  $A$ , there exists elementary matrices (or identity matrices)  $E_1, E_2, E_3$  and  $E_4$  such that

$$E_1 E_2 E_3 E_4 I = A$$

The matrices  $E_1, E_2, E_3, E_4$  either leaves determinant unchanged or change sign or scale the determinant by some factor. Correspondingly, elementary row operations leaves area unchanged or scales it. This gives  $\det A = \det E_1 \cdot \det E_2 \cdot \det E_3 \cdot \det E_4$  which is scaled area of rectangle with columns given by  $I$ .

## Result

2 of 2

Using elementary matrix operation, show that  $I$  can be scaled to  $A$ .

10. a

Suppose that  $(I_m - AB)$  is invertible. Consider the matrix

$$D = I_n + B(I_m - AB)^{-1}A$$

Then

$$\begin{aligned}
(I_n - BA)D &= (I_n - BA)(I_n + B(I_m - AB)^{-1}A) \\
&= I_n - BA + B(I_m - AB)^{-1}A - BAB(I_m - AB)^{-1}A \\
&= I_n - BA + B((I_m - AB)^{-1}A - AB(I_m - AB)^{-1}A) \\
&= I_n - BA + B((I_m - AB)^{-1} - AB(I_m - AB)^{-1})A \\
&= I_n - BA + B \underbrace{(I_m - AB)^{-1}(I_m - AB)}_I A \\
&= I_n - BA + BA \\
&= I_n
\end{aligned}$$

Here we used the Distributive Property multiple times.

Thus,  $D$  is an inverse of  $(I_n - BA)$ , so  $(I_n - BA)$  is truly invertible.

Now assume that  $(I_n - BA)$  is invertible. Define

$$C = I_m + A(I_n - BA)^{-1}B$$

Then, similarly to above,

$$\begin{aligned} (I_m - AB)C &= (I_m - AB)(I_m + A(I_n - BA)^{-1}B) \\ &= I_m - AB + A(I_n - BA)^{-1}B - ABA(I_n - BA)^{-1}B \\ &= I_m - AB + A((I_n - BA)^{-1}B - BA(I_n - BA)^{-1}B) \\ &= I_m - AB + A((I_n - BA)^{-1} - BA(I_n - BA)^{-1})B \\ &= I_m - AB + A \underbrace{(I_n - BA)^{-1}(I_n - BA)}_I B \\ &= I_m - AB + AB \\ &= I_m \end{aligned}$$

Thus,  $C$  is an inverse of  $(I_m - AB)$ , so  $(I_m - AB)$  is truly invertible.

### Result

Hint: if  $I_m - AB$  is invertible, consider the matrix  $D = I_n + B(I_m - AB)^{-1}A$ .

11. a

(a) The system of linear equations associated with this Dirichlet problem is

$$\begin{bmatrix} 1 & 0 & -1/4 & 0 & 0 \\ 0 & 1 & -1/4 & 0 & 0 \\ -1/4 & -1/4 & 1 & -1/4 & -1/4 \\ 0 & 0 & -1/4 & 1 & 0 \\ 0 & 0 & -1/4 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{-1,0} \\ x_{0,-1} \\ x_{0,0} \\ x_{0,1} \\ x_{1,0} \end{bmatrix} = \begin{bmatrix} 1/4 \\ 0 \\ 0 \\ 3/4 \\ 1/4 \end{bmatrix}$$

which has solution

$$\begin{bmatrix} x_{-1,0} \\ x_{0,-1} \\ x_{0,0} \\ x_{0,1} \\ x_{1,0} \end{bmatrix} = \begin{bmatrix} \frac{17}{48} \\ \frac{41}{48} \\ \frac{5}{12} \\ \frac{5}{48} \\ \frac{17}{48} \end{bmatrix}$$

- (b) By the Laplace equation each value  $f(u, v)$  of a harmonic function is the average of the values of its four neighbors. The average of a set of values is always lower or equal to the maximum of these values. The maximum principle then follows by induction on the maximum distance  $l_R$  between a point in  $R$  and a point in  $\partial R$ .

If  $l_R=1$  each point of  $R$  is surrounded by boundary points so its value by  $f$  is at most a value of the boundary. By induction assuming this holds for  $l_R = k$  if the maximum distance is  $k+1$  then considering  $R' \subset R$  the subset of points at a distance  $k$  from the boundary then by the inductive hypothesis the maximum principle for the region  $R'$  holds (considering the points  $R - R'$  as being part of the boundary now), so the maximum value occurs in either the boundary or at the points in  $R - R'$ . But then the maximum principle for  $R - R'$  (note that the boundary of  $R - R'$  is contained in  $R'$ ) implies the values at the points in  $R - R'$  are at most the values of points in  $R'$ . Therefore the points with maximum value must occur at the boundary of  $R$ .

- (c) The existence of a unique solution is equivalent to  $L$  being an invertible matrix. Now note that  $L$  is of the form  $I - A$  with

$$A = [a_{xy}^{uv}] = \begin{bmatrix} 1/4, & |(u, v) - (x, y)| = 1 \\ 0, & |(u, v) - (x, y)| \neq 1 \end{bmatrix}$$

Therefore  $L^{-1} = \sum_{\mathbb{N}} A^n$ .

## Result

- (a) The system of linear equations associated with this Dirichlet problem is

$$\begin{bmatrix} 1 & 0 & -1/4 & 0 & 0 \\ 0 & 1 & -1/4 & 0 & 0 \\ -1/4 & -1/4 & 1 & -1/4 & -1/4 \\ 0 & 0 & -1/4 & 1 & 0 \\ 0 & 0 & -1/4 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{-1,0} \\ x_{0,-1} \\ x_{0,0} \\ x_{0,1} \\ x_{1,0} \end{bmatrix} = \begin{bmatrix} 1/4 \\ 0 \\ 0 \\ 3/4 \\ 1/4 \end{bmatrix}$$

which has solution

$$\begin{bmatrix} x_{-1,0} \\ x_{0,-1} \\ x_{0,0} \\ x_{0,1} \\ x_{1,0} \end{bmatrix} = \begin{bmatrix} \frac{17}{48} \\ \frac{41}{48} \\ \frac{5}{12} \\ \frac{5}{48} \\ \frac{17}{48} \end{bmatrix}$$

- (b) The maximum principle follows by induction on the maximum distance  $l_R$  between a point in  $R$  and a point in  $\partial R$ .
- (c) Since  $L$  is of the form  $I - A$  we have  $L^{-1} = \sum_{\mathbb{N}} A^n$ .

# 2

---

## Chapter 2

### Section 1

1. a

Take  $a, b, c \in S$ . Then

$$(ab)c = ac = a$$

On the other hand,

$$a(bc) = ab = a$$

Thus,

$$(ab)c = a = a(bc),$$

so

$$(ab)c = a(bc),$$

and this law is indeed associative.

Now suppose that  $e \in S$  is the identity. Then, for each  $a$  in  $S$ , we must have

$$ae = ea = a$$

However,

$$ea = a$$

yields, since  $ea = e$  by definition of this law of composition,

$$e = a$$

So, the identity is equal to all elements of  $S$ ! This is clearly possible if and only if  $S = \{e\}$ , that is,  $S$  has only one element.

#### Result

Prove that  $(ab)c = a(bc)$  for all  $a, b, c \in S$ .

For the second part, prove that  $S$  must contain only one element.

2. a

(a)

Let  $a$  be such that  $l, r$  with the properties  $la = 1$  and  $ar = 1$  exist. Multiply the equality  $la = 1$  by  $r$  from the right:

$$(la)r = 1r$$

Since  $l$  is identity,  $1r = r$ . On the other hand, since the law is associative, and  $r$  is the right inverse,

$$(la)r = l(ar) = l1 = l$$

Thus,

$$la = 1$$

after multiplying by  $r$  from the right yields

$$l = r$$

as required. The other two statements now follow directly:

$$la = ra = 1$$

$$ar = 1$$

Thus,  $a^{-1} = r (= l)$ .

(b)

Suppose that  $b_1, b_2$  are inverses of  $a$ . We must prove that  $b_1 = b_2$ . Start:

$$b_1 = b_1 1 = b_1(ab_2) = (b_1a)b_2 = 1b_2 = b_2$$

In the first equality we used that  $1$  is the identity, in the second we used that  $b_2$  is an inverse of  $a$ , then we used the associative property, then the fact that  $b_1$  is an inverse of  $a$ , and finally that  $1$  is the identity again.

Thus,

$$b_1 = b_2,$$

so the inverse is unique.

(c)

Using the associativity and the definitions of inverse and identity multiple times:

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a1a^{-1} = aa^{-1} = 1$$

$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}1b = b^{-1}b = 1$$

Thus,  $b^{-1}a^{-1}$  is the (unique!) inverse of  $ab$ , so

$$(ab)^{-1} = b^{-1}a^{-1}$$

---

## Result

3 c

Hints:

For the first part, multiply the equality  $la = 1$  by  $r$  from the right.

For the second part: Suppose that  $b_1, b_2$  are inverses of  $a$  and prove that we must have  $b_1 = b_2$ .

For the last part, just prove it by definition of the inverse.



### 3. a

Intuitively, both "discrepancies" happen because 1 is not in the range of  $s$ ; that is, there exists no  $n \in \mathbb{N}$  such that  $s(n) = 1$ .

Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be any function from the set of natural numbers to itself. Let  $m = f(1)$ . Then

$$(s \circ f)(1) = s(f(1)) = s(m) = m + 1 > 1$$

Thus,  $s \circ f$  cannot be the identity function, since then we would have  $(s \circ f)(1) = 1$ . Since  $f$  was taken arbitrarily, then  $s$  has no right inverse.

Now let  $g_m$ , where  $m \in \mathbb{N}$ , be a function from a set of natural numbers to itself defined as follows:

$$g_m(n) = \begin{cases} m & \text{for } n = 1 \\ n - 1 & \text{for } n > 1 \end{cases}$$

Now, for every  $n \in \mathbb{N}$ ,

$$(g_m \circ s)(n) = g_m(s(n)) = g_m(\underbrace{n+1}_{>1}) = n$$

Thus,  $g_m \circ s$  is an identity function, so  $g_m$  is a left inverse of  $s$ . Since  $m$  was an arbitrary taken natural number, we conclude that infinitely many left inverses of  $s$  exist.

#### Result

Hint: both "discrepancies" occur because 1 is not in the range of  $s$ .

## Section 2

### 1. a



We first need to write all elements of  $S_3$ . We will write them as products of disjoint cycles (including the cycles of 1 element):

$$p_1 = (1)(2)(3)$$

$$p_2 = (12)(3)$$

$$p_3 = (1)(23)$$

$$p_4 = (13)(2)$$

$$p_5 = (123)$$

$$p_6 = (132)$$

Now we need to find all products. For example,

$$p_3 \circ p_5 = [(1)(23)](123) = (321) = (132) = p_6,$$

since  $p_5$  sends 1 to 2 and  $p_3$  sends 2 to 3, and so on. We also used the fact that the cycle can be written with any element in it written on the first place; we just have to make sure we do not change the positions between elements.

So, the table will be as follows:

$\circ$	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$
$p_1$	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$
$p_2$	$p_2$	$p_1$	$p_5$	$p_6$	$p_3$	$p_4$
$p_3$	$p_3$	$p_6$	$p_1$	$p_5$	$p_4$	$p_2$
$p_4$	$p_4$	$p_5$	$p_6$	$p_1$	$p_2$	$p_3$
$p_5$	$p_5$	$p_4$	$p_2$	$p_3$	$p_6$	$p_1$
$p_6$	$p_6$	$p_3$	$p_4$	$p_2$	$p_1$	$p_5$

## Result

$$p_1 = (1)(2)(3)$$

$$p_2 = (12)(3)$$

$$p_3 = (1)(23)$$

$$p_4 = (13)(2)$$

$$p_5 = (123)$$

$$p_6 = (132)$$

$\circ$	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$
$p_1$	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$
$p_2$	$p_2$	$p_1$	$p_5$	$p_6$	$p_3$	$p_4$
$p_3$	$p_3$	$p_6$	$p_1$	$p_5$	$p_4$	$p_2$
$p_4$	$p_4$	$p_5$	$p_6$	$p_1$	$p_2$	$p_3$
$p_5$	$p_5$	$p_4$	$p_2$	$p_3$	$p_6$	$p_1$
$p_6$	$p_6$	$p_3$	$p_4$	$p_2$	$p_1$	$p_5$

2. a

Let  $T$  be a set of all invertible elements of  $S$ .

We first must prove that for all  $a, b \in T$ , we have  $ab \in T$  (so for the law of composition to be well-defined on  $T$ ). However, the product of two invertible elements is invertible, so  $ab$  truly is in  $T$ .

Now for the associativity. However, this follows directly, because the law of composition is associative on  $S$ , and  $T$  is a subset of  $S$ .

Moreover,  $S$  has an identity; we will denote it by  $e$ . However,

$$ee = e,$$

so  $e$  is invertible (inverse being  $e$  itself)! Thus,  $e \in T$ , and  $T$  has an identity element.

Finally, let  $a \in T$ . Then there exists an inverse  $a^{-1}$ . For now we know only that  $a^{-1} \in S$ ; we must prove that it is also an element of  $T$ ! However,

$$aa^{-1} = a^{-1}a = e$$

actually means that  $a$  is an inverse of  $a^{-1}$ ; hence,  $a^{-1}$  is invertible, and  $a^{-1} \in T$ . Thus, every element of  $T$  has an inverse in  $T$ .

## Result

2 of 2

Denote by  $T$  the set of invertible elements of  $S$ . Prove that a restriction of the law of composition on  $T$  is well-defined, that it is associative, that  $T$  has an identity, and that each element of  $T$  has an inverse in  $T$ .

3. a

(a)

First multiply by  $x^{-1}$  from the left:

$$yz^{-1}w = x^{-1}$$

Now multiply by  $w^{-1}$  from the right:

$$yz^{-1} = x^{-1}w^{-1}$$

Finally, multiply by  $z$  from the right:

$$y = x^{-1}w^{-1}z$$

(b)

Start from  $xyz = 1$ . Multiply by  $z^{-1}$  from the right, and then by  $y^{-1}$  from the right to get

$$x = z^{-1}y^{-1} = (yz)^{-1}$$

Now multiply by  $yz$  from the left:

$$yzx = 1$$

For the second statement, let  $G = S_3$ , let

$$x = (1)(23), \quad y = (12)(3), \quad z = (123)$$

Then  $xy = (132)$  and  $xyz = (132)(123) = (1)(2)(3)$ , which is an identity in  $S_3$ . On the other hand,  $yx = (123)$  and  $yxz = (123)(123) = (132)$ , which is not an identity. Thus, the second statement does not hold.

## Result

24

$$(a) \quad y = x^{-1}w^{-1}z$$

(b) The first statement holds, the second does not.

4. a

(a)

**Subset?** Since every matrix with real entries can be interpreted as a matrix with complex entries, we conclude that  $H \subseteq G$ .

**Closure?** For every  $A, B \in H$  we have  $AB \in H$  since  $H$  is a group itself.

**Identity?** The identity in  $G$  is  $I_n$ , the identity  $n \times n$  matrix. Since every entry of  $I_n$  is real (every entry of  $I_n$  is either 1 or 0), we conclude that  $I_n \in H$ . Thus,  $H$  contains the identity.

**Inverse?** For every  $A \in H$  we have  $A^{-1} \in H$  since  $H$  is a group itself.

Thus,  $H$  is a subgroup of  $G$ .

(b)

**Subset?** Since  $\mathbb{R}^\times$  is a set of nonzero real numbers, we conclude that  $H \subseteq G$ .

**Closure?** For every  $x, y \in H$  we have  $xy \in H$ ; this is seen by looking at all possible combinations of  $x$  and  $y$ :

$$x = 1, y = 1 \Rightarrow xy = 1 \in H$$

$$x = -1, y = 1 \Rightarrow xy = -1 \in H$$

$$x = 1, y = -1 \Rightarrow xy = -1 \in H$$

$$x = -1, y = -1 \Rightarrow xy = 1 \in H$$

**Identity?** The identity in  $G$  is 1, which is also in  $H$ .

**Inverse?** For every  $x \in H$  we have  $x^{-1} \in H$ ; again, this is seen by looking at all cases with respect to  $x$ :

$$x = 1 \Rightarrow 1 \cdot 1 = 1 \Rightarrow x^{-1} = 1 \in H$$

$$x = -1 \Rightarrow -1 \cdot (-1) = 1 \Rightarrow x^{-1} = -1 \in H$$

Thus,  $H$  is a subgroup of  $G$ .

(c)

Notice that 0 is an identity in  $\mathbb{Z}^+$ , but  $0 \notin H$ , so  $H$  is **not** a subgroup of  $G$ .

(d)

Notice that  $H$  is not even a subset of  $G$  (matrices in  $H$  have determinant 0), so  $H$  is clearly **not** a subgroup of  $G$ .

---

### Result

3 of 3

$H$  is a subgroup of  $G$  in (a) and (b). It is not a subgroup in (c) and (d).

5. a

Let  $1_G$  be the identity in  $G$ , and let  $1_H$  be the identity in  $H$ . We must prove that  $1_G = 1_H$ .

First of all, since  $H \subseteq G$ ,  $1_H \in G$ . This means that, since  $G$  is a group,  $1_H^{-1}$  exists, and

$$1_H 1_H^{-1} = 1_G$$

Now notice that  $1_H^{-1} \in H$  since  $1_H \in H$  and  $H$  is a subgroup of  $G$ . Now we can apply the closure property to conclude that  $1_G = 1_H 1_H^{-1} \in H$ .

Finally,

$$1_H = 1_H 1_G = 1_G,$$

where in the first equality we used the fact that  $1_H \in G$  and that  $1_G$  is the identity in  $G$ , while in the second equality we used the fact that  $1_G \in H$  and that  $1_H$  is the identity in  $H$ .

---

### Result

2 of 2

Let  $1_G$  be the identity in  $G$ ,  $1_H$  be the identity in  $H$ . First prove that  $1_G \in H$ . Then, argue why the following holds:

$$1_H = 1_H 1_G = 1_G$$

6. a

### Associativity?

Let  $x, y, z \in G^\circ$ . Then

$$x * (y * z) = (y * z)x = (zy)x$$

and

$$(x * y) * z = z(x * y) = z(yx)$$

Since the associativity in  $G$  holds,

$$(zy)x = z(yx)$$

Thus,

$$x * (y * z) = (x * y) * z,$$

which means that the associativity in  $G^\circ$  holds.

### Identity?

Let  $1$  be the identity in  $G$ . Since  $G$  and  $G^\circ$  "share" the same set, we know that  $1 \in G^\circ$ . We will now prove that it is also the identity in  $G^\circ$ .

Let  $x \in G^\circ$ . Then

$$1 * x = x1 = x$$

$$x * 1 = 1x = x$$

since  $1$  is an identity in  $G$ . Therefore,

$$1 * x = x * 1 = x,$$

so  $1$  is truly the identity in  $G^\circ$ .

### Inverse?

Let  $x \in G^\circ$  be arbitrarily taken. Then, since  $G^\circ$  and  $G$  "share" the same set,  $x \in G$ . It also has an inverse  $x^{-1}$  in  $G$  since  $G$  is a group. We will prove that  $x^{-1}$  is an inverse of  $x$  in  $G^\circ$ . For this,

$$x * x^{-1} = x^{-1}x = 1$$

$$x^{-1} * x = xx^{-1} = 1,$$

which prove that  $x^{-1}$  is truly an inverse of  $x$  in  $G^\circ$ .

Conclusion. Since all properties of a group hold,  $G^\circ$  is a group.

---

### Result

Prove that the properties of a group hold.

## Section 3

1. a

We perform the Euclidean Algorithm:

$$\begin{aligned}312 &= 123 \cdot 2 + 66 \\123 &= 66 \cdot 1 + 57 \\66 &= 57 \cdot 1 + 9 \\57 &= 9 \cdot 6 + 3 \\9 &= 3 \cdot 3 + 0\end{aligned}$$

Therefore,

$$\gcd(123, 312) = \gcd(123, 66) = \gcd(66, 57) = \gcd(57, 9) = \gcd(9, 3) = \gcd(3, 0) = 3$$

because clearly  $\gcd(m, 0) = m$ , for all  $m \in \mathbf{Z}$ . Thus,  $d = 3$ .

Now we write the above equalities a bit differently:

$$\begin{aligned}312 &= 123 \cdot 2 + 66 \Rightarrow 66 = 312 - 123 \cdot 2 \\123 &= 66 \cdot 1 + 57 \Rightarrow 57 = 123 - 66 \\66 &= 57 \cdot 1 + 9 \Rightarrow 9 = 66 - 57 \\57 &= 9 \cdot 6 + 3 \Rightarrow 3 = 57 - 9 \cdot 6 \\9 &= 3 \cdot 3 + 0\end{aligned}$$

Therefore,

$$\begin{aligned}3 &= 57 - 9 \cdot 6 \\&= 57 - (66 - 57) \\&= -66 + 57 \cdot 2 \\&= -66 + (123 - 66) \cdot 2 \\&= 123 \cdot 2 - 66 \cdot 3 \\&= 123 \cdot 2 - (312 - 123 \cdot 2) \cdot 2 \\&= 312 \cdot (-2) + 123 \cdot 6\end{aligned}$$

Thus,

$$r = -2, s = 6$$

---

## Result

$$d = 3, r = -2, s = 6$$

2. a

First, since  $a$  and  $b$  are positive,  $p$  cannot divide  $a$  and  $p$  cannot divide  $b$  (since  $a < p$  and  $b < p$ ). Now let  $e$  be a (positive) common divisor of  $a$  and  $b$ ; that is,  $e$  divides  $a$  and  $e$  divides  $b$ . Then  $a = ke$  and  $b = le$  for some positive integers  $k, l$ , and

$$ke + le = p \Rightarrow (k + l)e = p$$

Thus,  $e$  divides  $p$ . Since  $p$  is prime, only 1 and  $p$  divide  $p$ , so  $e = 1$  or  $e = p$ . However,  $e = p$  is impossible since  $e$  divides  $a$  and  $b$ , but  $p$  does not.

Therefore,  $e = 1$ . By definition,  $d = \gcd(a, b)$  is a positive integer which divides both  $a$  and  $b$ . Since the only positive integer which divides both  $a$  and  $b$  is 1, we conclude that  $\boxed{d = 1}$ , as required.

## Result

2 of 2

Prove that the only positive integer which divides both  $a$  and  $b$  is 1.

3. a

(a)

We define it as a positive integer  $d$  such that

$$\mathbb{Z}d = \mathbb{Z}a_1 + \dots + \mathbb{Z}a_n \quad (1)$$

Such  $d$  exists because

$$S = \mathbb{Z}a_1 + \dots + \mathbb{Z}a_n = \{n \in \mathbb{Z} \mid n = r_1a_1 + \dots + r_na_n, r_1, \dots, r_n \in \mathbb{Z}\}$$

is a subgroup of  $\mathbb{Z}$ . Proof:

Subset? Clearly,  $S \subseteq \mathbb{Z}$ .

Closure? Let  $x, y \in S$ . Then

$$x = x_1a_1 + \dots + x_na_n, \quad y = y_1a_1 + \dots + y_na_n$$

and

$$x + y = (x_1 + y_1)a_1 + \dots + (x_n + y_n)a_n$$

which is in  $S$  because  $x_1 + y_1, \dots, x_n + y_n \in \mathbb{Z}$ .

Inverse? Let  $x \in S$ ,

$$x = x_1a_1 + \dots + x_na_n$$

Then its inverse is

$$-x = (-x_1)a_1 + \dots + (-x_n)a_n,$$

which is also in  $S$ , since  $-x_1, \dots, -x_n \in \mathbb{Z}$ .

Thus,  $S$  is a subgroup of  $\mathbb{Z}$ , so there exists a positive integer  $d$  such that (1) holds because of Theorem 2.3.3.

Moreover,  $d \in S$  means that there exist  $r_1, \dots, r_n$  such that

$$d = r_1a_1 + \dots + r_na_n$$



(b)

Denote by

$$a'_1 = a_1/d, \quad \dots, \quad a'_n = a_n/d$$

Suppose that  $e$  is a positive integer which divides  $a'_1, \dots, a'_n$ . Then  $a'_i = ek_i$ , for  $i = 1, \dots, n$ , where  $k_i$  are positive integers. This also means that  $a_i = da'_i = (de)k_i$ . So,  $de$  divides all  $a_i$ . By the generalization of Proposition 2.3.5 (c), we conclude that  $de$  divides  $d$ . However,  $de \geq d$  since  $e \geq 1$ ; so, we must have that  $de = d$ , so  $e = 1$ .

Similarly, we can conclude that  $d' = \gcd(a'_1, \dots, a'_n)$  divides all  $a'_i$ . This is because

$$\mathbb{Z}d' = \mathbb{Z}a'_1 + \dots + \mathbb{Z}a'_n$$

So,  $a'_i \in \mathbb{Z}d'$ , which means that  $a'_i = l_id'$  for some positive integer  $l_i$ . This in turn means that  $d'$  divides  $a'_i$ .

Finally, since the only positive integer which divides all  $a'_i$  is 1, we conclude that  $d' = 1$ .

### Result

(a) Just copy the definition of the gcd of two integers.

(b) Hint: prove that the only positive integer which divides  $a_1/d, \dots, a_n/d$  is 1.

## Section 4

1. a

Multiply the equality  $a^3b = ba^3$  by  $a^3$  from the left:

$$a^6b = a^3ba^3$$

Now use that  $a^3b = ba^3$  to conclude that

$$a^3ba^3 = (a^3b)a^3 = (ba^3)a^3 = ba^6$$

Therefore,

$$a^6b = ba^6$$

Now multiply this equality by  $a$  from the left:

$$a^7b = aba^6$$

Since  $a^7 = 1$ , we get

$$b = aba^6$$

Finally, multiply by  $a$  from the right:

$$ba = ab \Rightarrow ab = ba$$

### Result

2 of 2

Multiplying by  $a^3$  from the left one easily gets  $a^6b = ba^6$ . Then, multiply by  $a$  from the left and then from the right to get the desired equality.



2. a

(a)

Using De Moivre's formula, if  $z^n = 1$ , then

$$z = e^{2k\pi i/n},$$

where  $k = 0, 1, \dots, n-1$ . Thus, if we denote by  $S$  the set of  $n$ th roots of unity,

$$S = \{1, e^{2\pi i/n}, \dots, e^{2(n-1)\pi i/n}\}$$

We first prove that it is a subgroup of  $\mathbb{C}^\times$ :

Subset? First of all,  $S \subseteq \mathbb{C}^\times$  is trivial.

Closure? Let  $x, y \in S$ . Then  $x = e^{2k\pi i/n}$ ,  $y = e^{2l\pi i/n}$ , for some  $k, l \in \{0, 1, \dots, n-1\}$ . Now,

$$xy = e^{2k\pi i/n} e^{2l\pi i/n} = e^{2(k+l)\pi i/n}$$

Moreover, notice that, if  $k+l = an+b$ , with  $a$  and  $b$  positive integers,  $b \in \{0, 1, \dots, n-1\}$ , then

$$xy = e^{2(an+b)\pi i/n} = \underbrace{e^{2a\pi}}_1 \cdot e^{2b\pi i/n} = e^{2b\pi i/n} \in S$$

Thus,  $xy \in S$ , so  $S$  is closed.

Inverse? Let  $x \in S$ . If  $x = 1$ , then  $x^{-1} = 1$ .

Otherwise,  $x = e^{2k\pi i/n}$  with  $k \in \{1, \dots, n-1\}$ . Then  $y = e^{2(n-k)\pi i/n} \in S$  (since  $n-k \in \{1, \dots, n-1\}$ ), and

$$xy = yx = e^{2n\pi i/n} = e^{2\pi i} = 1$$

Thus,  $y = x^{-1}$ , so  $x^{-1} \in S$ .

This proves that  $S$  truly is a subgroup of  $\mathbb{C}^\times$ .

To prove that it is cyclic, just notice that

$$\left(e^{2\pi i/n}\right)^k = e^{2k\pi i/n},$$

so  $S = \left\langle e^{2\pi i/n} \right\rangle$ .

**(b)**

The product is

$$\begin{aligned} 1 \cdot e^{2i\pi/n} \dots e^{2(n-1)i\pi/n} &= \prod_{k=1}^{n-1} e^{2ki\pi/n} \\ &= e^{2i\pi/n \cdot \sum_{k=1}^{n-1} k} \\ &= e^{2i\pi/n \cdot \frac{(n-1)n}{2}} \\ &= e^{(n-1)i\pi} \end{aligned}$$

Now, if  $n$  is odd, then  $n - 1$  is even, and

$$e^{(n-1)i\pi} = 1,$$

since  $e^{ix\pi}$  is periodic with period of 2 and it is 1 when  $x = 0$ .

Similarly, if  $n$  is even, then  $n - 1$  is odd, and

$$e^{(n-1)i\pi} = e^{i\pi} = -1$$

Thus, we can write

$$e^{(n-1)i\pi} = (-1)^{n-1}$$

So, the desired product is  $\boxed{(-1)^{n-1}}$ .

### Result

**(a)** Prove that  $S = \langle e^{2i\pi/n} \rangle$ .

**(b)**  $(-1)^{n-1}$

3. a

We first look at the case when  $ab$  has a finite order.

Assume that  $ab$  has an order  $n$ . Then  $(ab)^n = 1$ . So,

$$\underbrace{(ab)(ab) \cdots (ab)}_{n \text{ times}} = 1$$

Multiplying by  $a^{-1}$  from the left and by  $b^{-1}$  from the right,

$$\underbrace{(ba)(ba) \cdots (ba)}_{(n-1) \text{ times}} = a^{-1}b^{-1}$$

Now multiply by  $ba$  from the right:

$$(ba)^n = a^{-1}b^{-1}ba = a^{-1}a = 1$$

Thus,  $ba$  has a finite order, denote it by  $m$ , and we conclude that  $m \leq n$ , since  $m$  is the smallest positive integer such that  $(ba)^m = 1$ .

Similarly,

$$\begin{aligned} (ba)^m &= 1 \\ (ab)^{m-1} &= b^{-1}a^{-1} \\ (ab)^m &= b^{-1}a^{-1}ab = b^{-1}b = 1 \end{aligned}$$

Thus,  $n \leq m$ , from which we conclude  $n = m$ , as required.

Now suppose that  $ab$  has an infinite order. We must prove that  $ba$  also has an infinite order. Suppose that  $ba$  has a finite order  $m$ , then the previous proof shows that  $(ab)^m = 1$ , which is a contradiction.

## Result

Hint: look at cases when  $ab$  has a finite and infinite order independently.

### 4. a

The easiest example is  $G = \{1\}$ . Now we suppose that  $G \neq \{1\}$ .

First suppose that  $G$  is not cyclic. Then there exist nontrivial  $a, b \in G$  such that

$$b \notin \langle a \rangle$$

However, now  $\langle a \rangle$  is a proper subgroup of  $G$  (since it is not equal to neither  $\{1\}$  nor  $G$ ), which is a contradiction.

Thus,  $G$  must be cyclic, that is,

$$G = \langle a \rangle$$

If the order  $G$  is infinite, then, for example,

$$\langle a^2 \rangle$$

is clearly a proper subgroup of  $G$ . Thus, the order of  $G$  must be finite, say,  $n$ .

Moreover, now  $\langle a^k \rangle$  is a subgroup of  $G$  for every  $k \in \{0, 1, 2, \dots, n-1\}$  (we do not need to take into account other  $k$  because  $a^0 = a^n = 1$ , so we can easily "translate" every  $a^l$  to be of the form  $a^k$  with  $k \in \{0, 1, 2, \dots, n-1\}$ ). Moreover, the order of  $\langle a^k \rangle$  is equal to  $n/d$ , where  $d = \gcd(n, k)$  by Proposition 2.4.3. Thus,  $\langle a^k \rangle$  has an order strictly between 1 and  $n$  if and only if  $1 < d < n$ . Also, in this case,  $\langle a^k \rangle$  is a proper subgroup of  $G$  (because of the number of elements).

If  $n$  is not prime, then there exists  $k \in \{2, 3, \dots, n-1\}$  such that  $d = \gcd(k, n) > 1$  and  $d < n$  (for example, we can pick one  $k$  which divides  $n$ ; such exists because  $n$  is not prime). By above discussion,  $\langle a^k \rangle$  is then a proper subgroup of  $G$ . Thus,  $G$  must be of prime order.

So, it is necessary for  $G$  to be of prime order. Is it sufficient? In other words, does every cyclic group  $G$  of prime order satisfy the given condition?

Let  $G$  be a cyclic group of prime order. Let  $H$  be its subgroup. We will prove that  $H$  is also cyclic. If  $H = \{1\}$ , the result is trivial, so suppose  $H \neq \{1\}$ . Define

$$S = \{k \in \{1, 2, \dots, n-1\} \mid a^k \in H\}$$

Let  $k_0 = \min S$  (such exists because  $S \subseteq \{1, 2, \dots, n-1\}$ ). We will prove that

$$H = \langle a^{k_0} \rangle$$

First of all,  $\langle a^{k_0} \rangle \subseteq H$  is trivial, since  $a^{k_0} \in H$ , and  $H$  is a group (so  $a^{mk_0} \in H$  for every  $m \in \mathbb{Z}$ ).

Now let  $x \in H$ . Then, since  $H \subseteq G$ , and  $G = \langle a \rangle$ ,  $x = a^m$  for some  $m$ . Now we write  $m = qk_0 + r$ , where  $0 \leq r < k_0$ .

Moreover,

$$a^m = a^{qk_0+r} = (a^{k_0})^q a^r$$

Since  $a^m \in H$  and  $(a^{k_0})^q \in H$ , we conclude that

$$a^r = a^m \left( (a^{k_0})^q \right)^{-1} \in H$$

Furthermore,  $r = 0$ . Otherwise, if  $r > 0$ , then  $r \in S$  and  $r < k_0 = \min S$ , which is a contradiction.

Thus,

$$a^m = (a^{k_0})^q \in \langle a^{k_0} \rangle$$

Finally, we can complete this exercise. Every subgroup of  $G$  is of the form  $\langle a^k \rangle$ . Since  $n$  is prime,  $\gcd(k, n)$  is either 1 or  $n$ , so every subgroup of  $G$  is either  $G$  or  $\{1\}$ . This completes the proof.

## Result

3 of 3

$G = \{1\}$  or  $G$  is a cyclic group of prime order. (Hints. What if it is not cyclic? What if it is infinite and cyclic? What if its order is not prime?)

5. a

Let  $G = \langle a \rangle$ . Let  $H$  be some subgroup of  $G$ . Define

$$S = \{k \in \mathbb{Z} \mid a^k \in H\}$$

We will first prove that  $S$  is a subgroup of  $\mathbb{Z}^+$ .

Subset? Clearly  $S \subseteq \mathbb{Z}^+$ .

Closure? Let  $k, l \in S$ . Then  $a^k \in H$  and  $a^l \in H$ . Since  $H$  is a subgroup of  $G$ , it is closed, so  $a^{k+l} = a^k \cdot a^l \in H$ . Thus,  $k + l \in S$ .

Inverse? Let  $k \in S$ . Then  $a^k \in H$ . Since  $H$  is a subgroup of  $G$ ,  $a^{-k} = (a^k)^{-1} \in H$ . Thus,  $-k \in S$ .

Conclusion. Now we know that  $S$  is a subgroup of  $\mathbb{Z}^+$ . If  $S$  is trivial ( $S = \{0\}$ ), then  $H = \{1\}$ , which is trivially cyclic. If  $S$  is nontrivial, then by Theorem 2.3.3. there exists some positive integer  $b$  such that

$$S = \mathbb{Z}b$$

So,  $k \in S$  if and only if there exists some integer  $l$  such that  $k = lb$ . Moreover, this means that

$$H = \{a^{lb} \mid l \in \mathbb{Z}\}$$

by definition of  $S$ . Thus,

$$H = \langle a^b \rangle$$

which proves that  $H$  is cyclic.

## Result

2 of 2

Define  $S = \{k \in \mathbb{Z} \mid a^k \in H\}$ . Prove that  $S$  is a subgroup of  $\mathbb{Z}^+$ . Conclude the rest by using Theorem 2.3.3.

6. a

(a)

For the first case,

$$G = \{1, a, a^2, a^3, a^4, a^5\}$$

By Proposition 2.4.3., the order of  $\langle a^k \rangle$  is  $6/d$ , where  $d = \gcd(6, d)$ . So, the order of  $\langle a \rangle$  and  $\langle a^5 \rangle$  is 6, while for the others, the order is less than 6. Thus, only  $a$  and  $a^5$  generate the entire  $G$ .

For the second case,

$$G = \{1, a, a^2, a^3, a^4\}$$

Now notice that every element except 1 can generate the entire  $G$ , since  $d = 1$  for  $k = 1, 2, 3, 4$ .

For the third case,

$$G = \{1, a, a^2, a^3, a^4, a^5, a^6, a^7\}$$

Similarly, only  $a, a^3, a^5, a^7$  can generate the entire  $G$ .

(b)

From before,  $a^k$  can generate the entire  $G$  if and only if  $\gcd(k, n) = 1$ . Thus, the number of elements of  $G$  which can generate it is equal to the number of integers  $1 \leq k < n$  which are relatively prime to  $n$ .

(a) 2; 4; 4

(b) The number of elements of  $G$  which can generate it is equal to the number of integers  $1 \leq k < n$  which are relatively prime to  $n$ .

7. a

We will first prove that

$$xy = yx$$

This follows because

$$(xy)^2 = 1 \Rightarrow (xy)(xy) = 1$$

Now multiply by  $y^{-1}x^{-1}$  from the right:

$$xy = y^{-1}x^{-1}$$

However, since

$$y^2 = yy = 1, \quad x^2 = xx = 1,$$

we conclude that  $y^{-1} = y$  and  $x^{-1} = x$ . Thus,

$$xy = yx$$

as required.

Subset?  $1, x, y \in G$  is trivial. Furthermore,  $xy \in G$  since  $G$  is a group. Thus,  $H \subseteq G$ .

Closure? Let  $a, b \in H$ . We prove this by looking at all possibilities with respect to  $a, b$ .

$$\begin{aligned} a = 1, b = 1 &\implies ab = 1 \in H \\ a = 1, b = x &\implies ab = x \in H \\ a = 1, b = y &\implies ab = y \in H \\ a = 1, b = xy &\implies ab = xy \in H \\ a = x, b = 1 &\implies ab = x \in H \\ a = x, b = x &\implies ab = x^2 = 1 \in H \\ a = x, b = y &\implies ab = xy \in H \\ a = x, b = xy &\implies ab = x^2y = y \in H \\ a = y, b = 1 &\implies ab = y \in H \\ a = y, b = x &\implies ab = yx = xy \in H \\ a = y, b = y &\implies ab = y^2 = 1 \in H \\ a = y, b = xy &\implies ab = yxy = xy^2 = x \in H \\ a = xy, b = 1 &\implies ab = xy \in H \\ a = xy, b = x &\implies ab = xyx = yx^2 = y \in H \\ a = xy, b = y &\implies ab = xy^2 = x \in H \\ a = xy, b = xy &\implies ab = (xy)^2 = 1 \in H \end{aligned}$$

Inverse? It is enough to notice that, for every  $a \in H$ , we have  $a^2 = 1$ , so  $a^{-1} = a \in H$ .



Conclusion and order. Thus,  $H$  is a subgroup of  $G$ . To prove that it is of order 4, we have to prove that no two elements of  $H$  are equal.

Since  $x, y, xy$  are of order 2, they are certainly not equal to 1. Moreover,  $x \neq y$ , since if  $x = y$ , then  $xy = x^2 = 1$ , which is a contradiction with  $xy$  having order 2.

Now suppose that  $x = xy$ . Multiply by  $x$  from the left:

$$x = xy \Leftrightarrow x^2 = x^2y \Leftrightarrow y = 1,$$

since  $x^2 = 1$ . However, this is a contradiction since  $y \neq 1$ . Thus,  $x \neq xy$ .

Similarly, suppose that  $y = xy$ . Multiply by  $y$  from the right:

$$y = xy \Leftrightarrow y^2 = xy^2 \Leftrightarrow x = 1,$$

since  $y^2 = 1$ . However, this is again a contradiction since  $x \neq 1$ . Thus,  $y \neq xy$ .

So, we concluded that no two elements of  $H$  are equal, meaning that the order of  $H$  truly is 4.

## Result

4 of 4

First prove that  $xy = yx$ . Then prove that  $H$  is a subgroup of  $G$ . Finally, to prove that it has an order 4, prove that no two elements of  $H$  are equal.

8. a

(a)

Let  $A \in \text{GL}_n(\mathbb{R})$ . It is sufficient to prove that

$$E_m E_{m-1} \cdots E_1 A F_1 F_2 \cdots F_k = I_n$$

for elementary matrices of Type 1 and 3  $E_i, F_j$ . Then

$$A = E_1^{-1} \cdots E_m^{-1} \cdot F_k^{-1} \cdots F_1^{-1}$$

Now we construct such matrices inductively. If  $a_{11} = 0$ , then there must be some  $a_{1j} \neq 0$  (or else we have a zero row in  $A$ , which means that  $\det A = 0$  and  $A \notin \text{GL}_n(\mathbb{R})$ ). Add the  $j$ th column to the first; we can do that by Type 1 elementary matrix. Now we can "scale" that element to 1 by a Type 3 matrix. Finally, by using Type 1 matrices, we can get all zeros in other entries of the first row and column. Thus, there exist some  $E_1, \dots, E_r, F_1, \dots, F_s$  Type 1 and 3 matrices such that

$$E_r E_{r-1} \cdots E_1 A F_1 F_2 \cdots F_s = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix},$$

where  $B$  is an  $(n-1) \times (n-1)$  matrix. Since  $\det A = \det(1) \cdot \det B = \det B$ , we must have  $\det B \neq 0$ , so  $B \in \text{GL}_{n-1}(\mathbb{R})$ . Now we proceed inductively.

The base case is when  $C \in \text{GL}_1(\mathbb{R})$ . Then  $C = [c]$ ,  $c \neq 0$ , so it can easily be transformed into  $C = I_1 = [1]$ .

Thus, by construction,

$$E_m E_{m-1} \cdots E_1 A F_1 F_2 \cdots F_k = I_n$$

which completes the proof by discussion from the start of this proof.

(b)

Let  $A \in \text{SL}_n(\mathbb{R})$ . Similarly to (a), we will prove that there exist elementary matrices of Type 1 such that

$$E_m E_{m-1} \cdots E_1 A F_1 F_2 \cdots F_k = I_n$$

First observe the case when  $A$  is a  $2 \times 2$  matrix. Then

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

If  $a = 0$ , then  $b \neq 0$  (otherwise,  $A$  has a zero column, so  $\det A = 0 \neq 1$ ). Thus, we can add the second row to the first and get that  $a_{11} \neq 0$  by a Type 1 matrix.

Now we can multiply the first row and add it to the second to get 1 in  $a_{21}$ . Then we can multiply the second row and add it to the first such that  $a_{11} = 1$  (this is done by two Type 1 matrices). Finally, we can, as in (a), get

$$E_r \cdots E_1 A F_1 \cdots F_s = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$$

where  $B$  is a  $1 \times 1$  matrix. However,  $\det B = \det A = 1$ , so  $B = [1]$ , which completes the proof when  $A$  is a  $2 \times 2$  matrix.

When  $A$  is a  $n \times n$  matrix, we actually proceed the same as above; we get

$$E_r \cdots E_1 A F_1 \cdots F_s = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$$

where  $\det B = 1$  and  $B$  is a  $(n-1) \times (n-1)$  matrix, so we proceed by induction.

---

## Result

Both parts can be solved by proving that

$$E_m \cdots E_1 A F_1 \cdots F_k = I_n$$

by some permissible matrices  $E_i$  and  $F_j$ .

9. a

Let  $p \in S_4$ . When we write  $p$  as a product of disjoint cycles, we get one of the following cases:

- $p$  has 4 1-cycles. So,  $p = 1$  (the identity in  $S_4$ ), which is clearly not of order 2.
- $p$  has 1 2-cycle and 2 1-cycles. Since disjoint cycles commute, clearly  $p^2 = 1$ . So, such  $p$  are of degree 2. There are  $\binom{4}{2} = 6$  such permutation (we choose 2 elements which will be in a 2-cycle; the other elements are in 1-cycles).
- $p$  has 2 2-cycles. Again, disjoint cycles commute, so  $p^2 = 1$ . So, such  $p$  are of degree 2. There are 3 such permutation, because we choose which of the other 3 elements will be in a cycle with the first element; the remaining two elements are in other cycle.
- $p$  has 1 3-cycle and a 1-cycle. This  $p$  is of degree 3, not 2.
- $p$  has a 4-cycle. This  $p$  is of degree 4, not 2.

To conclude, there are 9 such permutations.

## Result

2 of 2

10. a



For the first statement, we consider  $\text{GL}_2(\mathbb{R})$ , and

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1/2 \\ 2 & 0 \end{bmatrix}$$

(clearly  $\det A = \det B = -1 \neq 0$ , so  $A, B \in \text{GL}_2(\mathbb{R})$ ). Moreover, it is easy to check that

$$A^2 = B^2 = I_2,$$

so  $A$  and  $B$  are of degree 2.

Now look at their product:

$$AB = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

It is easy to see that

$$(AB)^n = \begin{bmatrix} 2^n & 0 \\ 0 & (1/2)^n \end{bmatrix}$$

for all  $n \in \mathbb{N}$ . Thus,  $(AB)^n \neq I_2$  for all  $n \in \mathbb{N}$ , meaning that  $AB$  is of infinite order.

On the other hand, if a group is commutative, then the product of elements of finite order is of finite order.

Suppose that  $G$  is abelian, and let  $x \in G$  be of order  $n$ ,  $y \in G$  be of order  $m$ . Then

$$(xy)^{nm} = \underbrace{(xy)(xy) \cdots (xy)}_{nm \text{ times}} \stackrel{(1)}{=} x^{nm} y^{nm} = (x^n)^m (y^m)^n = 1^m 1^n = 1,$$

where (1) holds because  $G$  is commutative.

## Result

Consider  $\text{GL}_2(\mathbb{R})$ , and

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1/2 \\ 2 & 0 \end{bmatrix}$$

If  $G$  is abelian, then the product of elements of finite order is of finite order.

11. a

**(a)**

We will prove that, for every  $p \in S_n$ , there exist transpositions  $\tau_1, \dots, \tau_r$  such that

$$\tau_r \cdots \tau_1 p = 1,$$

where  $1 \in S_n$  is an identity. We prove this by induction.

If  $n = 2$ , then  $S_2 = \{(1)(2), (12)\}$ . For  $p = (1)(2)$ , we can write

$$(12)(12)(1)(2) = 1 = (1)(2)$$

For  $p = (12)$ , we write

$$(12)(12) = (1)(2)$$

since  $(12)$  is a transposition. Thus, the statement holds for  $n = 2$ .

Now assume that the statement holds for  $2 \leq n < k$ , where  $k > 2$ . Let  $p \in S_k$ . Suppose that  $p(k) = i \neq k$  (if  $i = k$ , we skip the following step). Let  $\tau \in S_k$  by a transposition which swaps  $i$  and  $k$ . Then

$$(\tau \circ p)(k) = \tau(p(k)) = \tau(i) = k$$

Thus,  $\tau p$  is a permutation of  $S_k$  which we can interpret as a permutation of  $S_{k-1}$  (since  $k$  is fixed). Using induction,

$$\tau_r \cdots \tau_1(\tau p) = 1$$

Thus, the statement holds for  $n = k$ .

So, for every  $n \in \mathbb{N}$  and every  $p \in S_n$ , there exist transpositions such that

$$\tau_r \cdots \tau_1 p = 1$$

or

$$p = \tau_1^{-1} \cdots \tau_r^{-1} = \tau_1 \cdots \tau_r,$$

since transpositions are clearly of degree 2. Thus, transpositions generate  $S_n$ .

**(b)**

This is similar to above. Base case is  $n = 3$ . Let  $p \in A_3$ . Suppose that  $p(3) = i$ . Let  $\tau = (ji3)$ , where  $j \neq i, j \neq 3$ . Then

$$\tau \circ p(3) = \tau(p(3)) = \tau(i) = 3$$

Thus,  $\tau p$  is an even permutation (product of even permutations is even!), and it fixes one element. If it swaps 1 and 2, then it would be odd permutation. So,  $\tau p = 1$ , and we are done.

Now suppose that this holds for all  $3 \leq n < k$ , where  $k > 3$ . Let  $p \in S_k$ ,  $p(k) = i$ . Define  $\tau = (jik)$ , where  $j \neq i, j \neq k$ . Thus

$$\tau p(k) = k,$$

so  $\tau p \in A_{k-1}$ . Now apply induction:

$$\tau_r \cdots \tau_1 \tau p = 1$$

Finally, this decomposition yields

$$p = \tau^{-1} \tau_1^{-1} \cdots \tau_r^{-1},$$

where all permutations on the right side are clearly 3-cycles. (The inverse of  $(ijk)$  is  $(kji)$ .)

## Result

(a) Prove that there exist transpositions  $\tau_i$  such that

$$\tau_r \cdots \tau_1 p = 1,$$

where  $1 \in S_n$  is an identity.

(b) Same as in (a).

## Section 5

1. a

First let  $G = \langle a \rangle$ . Let  $x \in G'$ . Then there exists some  $g \in G$  such that  $\varphi(g) = x$ , since  $\varphi$  is surjective. However,  $G = \langle a \rangle$  and  $g \in G$  means that  $g = a^k$ , for some  $k \in \mathbb{Z}$ . Since  $\varphi$  is a homomorphism,

$$x = \varphi(g) = \varphi(a^k) = \varphi(\underbrace{a \cdots a}_{k \text{ times}}) = \underbrace{\varphi(a) \cdots \varphi(a)}_{k \text{ times}} = \varphi(a)^k \in \langle \varphi(a) \rangle$$

Therefore,  $G' \subseteq \langle \varphi(a) \rangle$ , so it is a subgroup of a cyclic group  $\langle \varphi(a) \rangle$ , and it also must be cyclic by Exercise 4.5.

Now let  $G$  be abelian, and let  $x, y \in G'$ . Then there exist some  $g, h \in G$  such that  $x = \varphi(g)$  and  $y = \varphi(h)$ . Now,

$$xy = \varphi(g)\varphi(h) \stackrel{(1)}{=} \varphi(gh) \stackrel{(2)}{=} \varphi(hg) \stackrel{(3)}{=} \varphi(h)\varphi(g) = yx$$

(in (1) and (3) we used that  $\varphi$  is a homomorphism and in (2) we used that  $G$  is abelian; that is, that  $gh = hg$ ).

So,  $xy = yx$  for all  $x, y \in G'$ , so  $G'$  is abelian.

## Result

Hint for the first part and  $G = \langle a \rangle$ ,  $x \in G'$ :

$$x = \varphi(g) = \varphi(a^k) = \varphi(\underbrace{a \cdots a}_{k \text{ times}}) = \underbrace{\varphi(a) \cdots \varphi(a)}_{k \text{ times}} = \varphi(a)^k \in \langle \varphi(a) \rangle$$

Hint for the second part:

$$xy = \varphi(g)\varphi(h) = \varphi(gh) = \varphi(hg) = \varphi(h)\varphi(g) = yx$$

where  $x, y \in G'$ .

2. a

$K \cap H$  is a subgroup of  $H$

We check the properties from the definition of a subgroup.

Subset? It is trivial that  $K \cap H \subseteq H$ .

Closure? Let  $x, y \in K \cap H$ . We must prove that  $xy \in K \cap H$ .

First of all,  $x, y \in K$ . Thus,  $xy \in K$  since  $K$  is a subgroup of  $G$ .

Similarly,  $x, y \in H$ , hence  $xy \in H$  since  $H$  is a subgroup of  $G$ .

Finally, since  $xy \in K$  and  $xy \in H$ , we conclude that  $xy \in K \cap H$ .

Inverse? Let  $x \in K \cap H$ . We must prove that  $x^{-1} \in K \cap H$ .

First of all,  $x \in K$ , so  $x^{-1} \in K$  since  $K$  is a subgroup of  $G$ .

Similarly,  $x \in H$ , so  $x^{-1} \in H$  since  $H$  is a subgroup of  $G$ .

Thus,  $x^{-1} \in K \cap H$ .

Conclusion. Now we can conclude that  $K \cap H$  is a subgroup of  $H$ .

$K \cap H$  is a normal subgroup of  $H$

Now suppose that  $K$  is a normal subgroup of  $G$ . To prove that  $K \cap H$  is a normal subgroup of  $H$ , let  $h \in H$  and  $k \in K \cap H$ , and we must prove that  $hkh^{-1} \in K \cap H$ .

First of all,  $k \in H$ , so  $hkh^{-1} \in H$  since  $H$  is a subgroup of  $G$ , so it is closed.

Now we want to use the fact that  $K$  is a normal subgroup of  $G$ . Since  $k \in K \cap H$ , then, specially,  $k \in K$ .

Moreover,  $h \in H$ , and since  $H$  is a subgroup of  $G$ , we know that  $h \in G$ . Similarly,  $h^{-1} \in G$ . Now we use the fact that  $K$  is a normal subgroup of  $G$  to conclude that  $hkh^{-1} \in K$ .

So,  $hkh^{-1} \in K$  and  $hkh^{-1} \in H$ , so  $hkh^{-1} \in K \cap H$ , which completes the proof.

**Result**

2 of 2

For the first part, check properties of the definition of a subgroup.

For the second part, prove that, for every  $h \in H$  and  $k \in K \cap H$ , we have that  $hkh^{-1} \in K \cap H$ .

3. a

First of all,  $a \neq 0$  since  $A$  is invertible, so  $\varphi$  is well-defined.

$\varphi$  is a homomorphism.

Let  $A_1 = \begin{bmatrix} a_1 & b_1 \\ 0 & d_1 \end{bmatrix}$ ,  $A_2 = \begin{bmatrix} a_2 & b_2 \\ 0 & d_2 \end{bmatrix}$ . Then

$$A_1 A_2 = \begin{bmatrix} a_1 a_2 & a_1 b_1 + b_1 d_2 \\ 0 & d_1 d_2 \end{bmatrix}$$

Thus,

$$\begin{aligned} \varphi(A_1 A_2) &= \varphi \left( \begin{bmatrix} a_1 a_2 & a_1 b_1 + b_1 d_2 \\ 0 & d_1 d_2 \end{bmatrix} \right) \\ &= (a_1 a_2)^2 \\ &= a_1^2 a_2^2 \\ &= \varphi \left( \begin{bmatrix} a_1 & b_1 \\ 0 & d_1 \end{bmatrix} \right) \varphi \left( \begin{bmatrix} a_2 & b_2 \\ 0 & d_2 \end{bmatrix} \right) \\ &= \varphi(A_1) \varphi(A_2) \end{aligned}$$

and  $\varphi$  is a homomorphism.

Kernel.

To determine its kernel, first recall that  $1$  is the identity in  $\mathbb{R}^\times$ . So, we need to find all  $A \in U$ ,

$$A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

such that

$$\varphi A = 1 \Leftrightarrow a^2 = 1 \Leftrightarrow \boxed{a = \pm 1}$$

So, the kernel of  $\varphi$  is the set of all invertible matrices  $A$  of the form

$$\begin{bmatrix} \pm 1 & b \\ 0 & d \end{bmatrix}$$

, or, written more precisely,

$$\ker \varphi = \left\{ A \in U \mid A = \begin{bmatrix} \pm 1 & b \\ 0 & d \end{bmatrix} \right\}$$

Image.

To determine its image, first notice that  $a^2 > 0$  for all  $a \in \mathbb{R} \setminus \{0\}$ , so  $\varphi(A) > 0$  for all  $A \in U$ . So,

$$\text{im} \varphi \subseteq \{x \in \mathbb{R} \mid x > 0\}$$

Now let  $x \in \mathbb{R}$ ,  $x > 0$  be arbitrarily taken. Then  $\sqrt{x} \in \mathbb{R}$ , so we can define a matrix

$$A = \begin{bmatrix} \sqrt{x} & 0 \\ 0 & 1 \end{bmatrix}$$

Notice that  $A \in U$  since it is invertible and upper-triangular. Moreover,

$$\varphi(A) = (\sqrt{x})^2 = x$$

Thus,  $x \in \text{im} \varphi$ , so

$$\text{im} \varphi = \{x \in \mathbb{R} \mid x > 0\}$$

To prove that  $\varphi$  is a homomorphism, prove that  $\varphi(A_1 A_2) = \varphi(A_1) \varphi(A_2)$  for all  $A_1, A_2 \in U$ .

Then,  $\ker \varphi$  consists of all  $A \in U$  such that  $a = \pm 1$ , and  $\operatorname{im} \varphi$  is the set of all positive real numbers.

4. a

Since  $|e^{ix}| = 1$ , for every  $x \in \mathbb{R}$ , we know that  $e^{ix} \neq 0$ , so  $f$  is well-defined.

*f* is a homomorphism.

Let  $x, y \in \mathbb{R}^+$ . Then

$$f(x+y) = e^{i(x+y)} = e^{ix} e^{iy} = f(x) f(y),$$

so  $f$  is truly a homomorphism.

Kernel.

Recall that 1 is the identity in  $\mathbb{C}^\times$ . So, we need to find all  $x \in \mathbb{R}^+$  such that

$$f(x) = 1 \Leftrightarrow e^{ix} = 1 \Leftrightarrow \cos x + i \sin x = 1 + i \cdot 0$$

This yields the system of equation of real numbers

$$\begin{aligned} \cos x &= 1 \\ \sin x &= 0 \end{aligned}$$

The solution of  $\cos x = 1$  is  $x = 2k\pi, k \in \mathbb{Z}$ .

The solution of  $\sin x = 0$  is  $x = k\pi, k \in \mathbb{Z}$ .

Thus, the solution of the system is  $x = 2k\pi, k \in \mathbb{Z}$ , so

$$\ker \varphi = \{x = 2k\pi, k \in \mathbb{Z}\}$$

Image.

Since  $|e^{ix}| = 1$  for all  $x \in \mathbb{R}$ , we conclude that  $\operatorname{im} \varphi \subseteq \{z \in \mathbb{C}^\times \mid |z| = 1\}$ . Now we want to prove that the converse inclusion holds.

Let  $z \in \mathbb{C}^\times, |z| = 1$ . Then  $z$  is located on a unit circle in complex plane. Thus, there exists a real number  $\mathbb{R}$  such that

$$z = \cos x + i \sin x = e^{ix}$$

Therefore,

$$f(x) = z$$

To conclude, we now have that

$$\operatorname{im} \varphi = \{z \in \mathbb{C}^\times \mid |z| = 1\}$$

## Result

First prove that  $f(x+y) = f(x)f(y)$  for all  $x, y \in \mathbb{R}^+$ .

Then,  $\ker \varphi = \{x = 2k\pi, k \in \mathbb{Z}\}$ ,  $\operatorname{im} \varphi = \{z \in \mathbb{C}^\times \mid |z| = 1\}$ .



5. a

[Subgroup?](#)

We check the properties from the definition of a subgroup.

[Subset?](#) First of all, clearly  $H \subseteq \text{GL}_n(\mathbb{R})$ , since  $\det M = \det A \det D \neq 0$ .

[Closure?](#) Let

$$M_1 = \begin{bmatrix} A_1 & B_1 \\ 0 & D_1 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} A_2 & B_2 \\ 0 & D_2 \end{bmatrix}$$

. Then

$$M_1 M_2 = \begin{bmatrix} A_1 A_2 & A_1 B_2 + B_1 D_2 \\ 0 & D_1 D_2 \end{bmatrix}$$

Since  $\det(A_1 A_2) = \det A_1 \det A_2 \neq 0$  and  $\det(D_1 D_2) = \det D_1 \det D_2 \neq 0$ ,  $A_1 A_2 \in \text{GL}_r(\mathbb{R})$  and  $D_1 D_2 \in \text{GL}_{n-r}(\mathbb{R})$ . Thus,  $M_1 M_2 \in H$ .

[Inverse?](#) Let

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

. Define

$$N = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{bmatrix}$$

( $A^{-1}$  and  $D^{-1}$  exist since  $A \in \text{GL}_r(\mathbb{R})$  and  $D \in \text{GL}_{n-r}(\mathbb{R})$ ). Also, clearly  $N \in H$ , since  $A^{-1} \in \text{GL}_r(\mathbb{R})$  and  $D^{-1} \in \text{GL}_{n-r}(\mathbb{R})$ .)

Then

$$\begin{aligned} MN &= \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{bmatrix} = \begin{bmatrix} I_r & -BD^{-1} + BD^{-1} \\ 0 & I_{n-r} \end{bmatrix} = I_n \\ NM &= \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{bmatrix} \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} I_r & A^{-1}B - A^{-1}B \\ 0 & I_{n-r} \end{bmatrix} = I_n \end{aligned}$$

Thus,  $N = M^{-1}$ , so  $M^{-1} \in H$ .

### Homomorphism and kernel.

Now denote by  $f$  the describe mapping. Let

$$M_1 = \begin{bmatrix} A_1 & B_1 \\ 0 & D_1 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} A_2 & B_2 \\ 0 & D_2 \end{bmatrix}$$

. Then

$$M_1 M_2 = \begin{bmatrix} A_1 A_2 & A_1 B_2 + B_1 D_2 \\ 0 & D_1 D_2 \end{bmatrix}$$

So,

$$f(M_1 M_2) = A_1 A_2 = f(M_1) f(M_2)$$

Thus,  $f$  is a homomorphism.

To find its kernel, first note that  $I_r$  is the identity in  $\text{GL}_r(\mathbb{R})$ . So we want to find all

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

, where  $A$  and  $D$  are invertible, such that

$$f(M) = I_r$$

However, by definition of  $f$ , it is clear that  $f(M) = I_r$  if and only if  $A = I_r$ . Thus,

$$\ker f = \left\{ \begin{bmatrix} I_r & B \\ 0 & D \end{bmatrix} \middle| D \in \text{GL}_{n-r}(\mathbb{R}) \right\}$$

### **Result**

3 of 3

To prove that  $H$  is a subgroup, check all properties from the definition of a subgroup.

To prove that the mapping, denoted by  $f$ , is a homomorphism, prove that  $f(M_1 M_2) = f(M_1) f(M_2)$  for all  $M_1, M_2 \in H$ . Moreover,

$$\ker f = \left\{ \begin{bmatrix} I_r & B \\ 0 & D \end{bmatrix} \middle| D \in \text{GL}_{n-r}(\mathbb{R}) \right\}$$

6. a



Let  $A$  be some matrix in the center. We will first prove that

$$AE_{ij} = E_{ij}A,$$

where  $E_{ij}$  is an  $n \times n$  matrix with zeros everywhere except on position  $(i, j)$  where we have 1.

First note that  $(I_n + E_{ij}) \in \text{GL}_n(\mathbb{R})$  because this is an elementary matrix. Thus,

$$A(I_n + E_{ij}) = (I_n + E_{ij})A$$

(that is,  $A$  commutes with it). Moreover,

$$A(I_n + E_{ij}) = A + AE_{ij}$$

and

$$(I_n + E_{ij})A = A + E_{ij}A$$

because of distributive property. Thus,

$$A + AE_{ij} = A + E_{ij}A$$

which yields

$$AE_{ij} = E_{ij}A$$

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Now consider  $E_{12}$  (it has a 1 on the intersection of the first row and second column). Then

$$\begin{aligned} AE_{12} &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0 & a_{11} & \dots & 0 \\ 0 & a_{21} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & \dots & 0 \end{bmatrix} \\ E_{12}A &= \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \end{aligned}$$

Since  $AE_{12} = E_{12}A$ , we conclude that  $a_{11} = a_{22}$  and that  $a_{j1} = 0$ , for  $j > 1$ , and  $a_{2j}$ , for  $j \neq 2$ .

Similarly, by multiplying by all  $E_{i,i+1}$ ,  $i = 1, 2, \dots, n - 1$ , we conclude that

$$a_{11} = a_{22} = a_{33} = \dots = a_{nn}$$

and

$$a_{ij} = 0, \quad i \neq j$$

Thus,

$$A = \begin{bmatrix} a & 0 & \dots & 0 \\ 0 & a & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a \end{bmatrix} = aI_n,$$

where  $a \neq 0$  since  $A \in \text{GL}_n(\mathbb{R})$ .

So, all  $A$  that are in the center must be of the above form. Now let  $A$  be of the above form, let  $B \in \text{GL}_n(\mathbb{R})$  be arbitrary. Then

$$AB = aI_n B = aB = Ba = B(aI_n) = BA$$

Thus, all such matrices commute.

To conclude, the center is

$$\{aI_n \mid a \in \mathbb{R} \setminus \{0\}\}$$

### Result

$$\{aI_n \mid a \in \mathbb{R} \setminus \{0\}\}$$

## Section 6

1. a

Denote by  $f$  the described mapping.

$f$  homomorphism?

Let  $x, y \in \mathbb{R}^+$ . Then

$$f(x+y) = \begin{bmatrix} 1 & x+y \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & y \\ 0 & 1 \end{bmatrix} = f(x)f(y)$$

(the second equality is checked by direct computation). Thus,  $f$  is truly a homomorphism.

$f$  injective?

Let  $x, y \in \mathbb{R}^+$  be such that  $f(x) = f(y)$ . Then

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & y \\ 0 & 1 \end{bmatrix},$$

so  $x = y$ . Thus,  $f$  is injective.

$f$  surjective?

Let

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$$

be arbitrarily taken matrix of this form. Then clearly  $f(a)$  is that matrix, so  $f$  is surjective.

Conclusion.

Since  $f$  is both injective and surjective, we conclude that it is also a bijection. Since it is also a homomorphism, it is an isomorphism.

---

## Result

2 of 2

It is an isomorphism. Prove that it is a homomorphism, and that it is injective and surjective.

2. a

Let  $\varphi(1) = m$ . First of all,

$$m = \varphi(1) = \varphi(1 + 0) = \varphi(1) + \varphi(0) = m + \varphi(0) \implies \varphi(0) = 0,$$

where the third equality follows from the fact that  $\varphi$  is a homomorphism. Moreover,

$$0 = \varphi(0) = \varphi(1 + (-1)) = \varphi(1) + \varphi(-1) = m + \varphi(-1) \implies \varphi(-1) = -m$$

Moreover, now we can get how  $\varphi$  operates on every  $k \in \mathbb{Z}$ !

For  $n \in \mathbb{N}$ , we have

$$\varphi(n) = \varphi(\underbrace{1 + 1 + \dots + 1}_{n \text{ times}}) = \underbrace{\varphi(1) + \dots + \varphi(1)}_{n \text{ times}} = \underbrace{m + \dots + m}_{n \text{ times}} = nm$$

and

$$\varphi(-n) = \varphi(\underbrace{(-1) + \dots + (-1)}_{n \text{ times}}) = \underbrace{\varphi(-1) + \dots + \varphi(-1)}_{n \text{ times}} = \underbrace{(-m) + \dots + (-m)}_{n \text{ times}} = -nm$$

Thus,

$$\varphi(k) = km$$

for every  $k \in \mathbb{Z}$ .

With this we described all homomorphisms  $\mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ .

If  $m = 0$ , then  $\varphi(k) = 0$  for all  $k \in \mathbb{Z}$ , so this  $\varphi$  is not injective.

Now let  $m \neq 0$ . Let  $k, l \in \mathbb{Z}$  be such that  $\varphi(k) = \varphi(l)$ . Then

$$km = lm \implies (k - l)m = 0$$

Thus,  $m = 0$  or  $k - l = 0$ . Since  $m \neq 0$ , we conclude that  $k - l = 0$ , so  $k = l$ , and  $\varphi$  is injective.

To find surjective homomorphisms, it is enough to notice that

$$\text{im}\varphi = \{\varphi(k) \mid k \in \mathbb{Z}\} = \{km \mid k \in \mathbb{Z}\} = \mathbb{Z}m$$

So,  $\varphi$  is surjective if and only if  $m = \pm 1$ .

Moreover, by combining all information above, we finally conclude that  $\varphi$  is an isomorphism if and only if  $m = \pm 1$ .

## Result

3 of 3

Prove that  $\varphi(k) = mk$ , for some  $m \in \mathbb{Z}$ . Then prove that  $\varphi$  is injective if and only if  $m \neq 0$ , and that it is surjective if and only if  $m = \pm 1$ . Then it easily follows that  $\varphi$  is an isomorphism if and only if  $m = \pm 1$ .

3. a

First notice that

$$(f \circ f)(x) = f(f(x)) = f\left(\frac{1}{x}\right) = x$$

and

$$(g \circ g)(x) = g\left(\frac{x-1}{x}\right) = \frac{\frac{x-1}{x} - 1}{\frac{x-1}{x}} = \frac{-\frac{1}{x}}{\frac{x-1}{x}} = -\frac{1}{x-1}$$

$$(g \circ g \circ g)(x) = g((g \circ g)(x)) = g\left(-\frac{1}{x-1}\right) = \frac{-\frac{1}{x-1} - 1}{-\frac{1}{x-1}} = \frac{1 + \frac{1}{x-1}}{\frac{1}{x-1}} = x$$

Thus,  $f$  is of degree 2, while  $g$  is of degree 3. Moreover,

$$(f \circ g)(x) = \frac{x}{x-1} = \frac{x-1+1}{x-1} = 1 + \frac{1}{x-1}$$

$$(f \circ g \circ g)(x) = f\left(-\frac{1}{x-1}\right) = -(x-1)$$

$$(g \circ f)(x) = g\left(\frac{1}{x}\right) = \frac{\frac{1}{x} - 1}{\frac{1}{x}} = \frac{\frac{1-x}{x}}{\frac{1}{x}} = 1-x = -(x-1)$$

$$(g \circ g \circ f)(x) = g(1-x) = \frac{1-x-1}{1-x} = \frac{x}{x-1} = 1 + \frac{1}{x-1}$$

$$(f \circ g \circ f)(x) = f(1-x) = \frac{1}{1-x} = -\frac{1}{x-1}$$

$$(g \circ f \circ g)(x) = g\left(\frac{x}{x-1}\right) = \frac{\frac{x}{x-1} - 1}{\frac{x}{x-1}} = \frac{\frac{x-(x-1)}{x-1}}{\frac{x}{x-1}} = \frac{1}{x}$$

Now we conclude that

$$H = \{\text{id}, f, g, g^2, fg, gf\}$$

is a group (notice that before we were actually checking all combinations of elements of  $H$  to confirm the closure property; the inverse property follows from the closure property and from the orders of  $f$  and  $g$ ).

Now, we write  $S_3$  as

$$S_3 = \{\text{id}, (12), (123), (123)^2, (12)(123), (123)(12)\},$$

since

$$(123)^2 = (132)$$

$$(12)(123) = (23)$$

$$(123)(12) = (13)$$

Now define the function  $\varphi : H \rightarrow S_3$  as

$$\varphi(\text{id}) = \text{id}$$

$$\varphi(f) = (12)$$

$$\varphi(g) = (123)$$

$$\varphi(g^2) = (123)^2$$

$$\varphi(fg) = (12)(123)$$

$$\varphi(gf) = (123)(12)$$

The fact that it is a homomorphism and that it is bijective follows directly from definition. Thus,

$$H \approx S_3,$$

as required.

## Result

3 of 3

Prove that

$$H = \{\text{id}, f, g, g^2, fg, gf\}$$

is a group generated by  $f$  and  $g$ . Prove that there are two elements of  $S_3$  that generate  $S_3$ . Use that to define an isomorphism.

## 4. a

We need to find some  $g \in G$  such that

$$ab = gbag^{-1}$$

Now notice that, for  $g = a$ ,

$$gbag^{-1} = abaa^{-1} = ab$$

So,  $ab$  and  $ba$  are conjugate.

## Result

Hint:  $g = a$  in the definition of conjugate elements.

## 5. a

We want to see if a matrix  $M \in \text{GL}_2(\mathbb{R})$ ,

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

such that

$$A = MBM^{-1},$$

exists. The above equality is equivalent to

$$AM = MB,$$

or

$$\begin{bmatrix} 3a & 3b \\ 2c & 2d \end{bmatrix} = \begin{bmatrix} a - 2b & a + 4b \\ c - 2d & c + 4d \end{bmatrix}$$

This yields the system of equations

$$\begin{aligned} 3a &= a - 2b \\ 3b &= a + 4b \\ 2c &= c - 2d \\ 2d &= c + 4d \end{aligned}$$

Simplifying,

$$\begin{aligned} a + b &= 0 \\ a + b &= 0 \\ c + 2d &= 0 \\ c + 2d &= 0 \end{aligned}$$

Thus,

$$\begin{aligned} a &= -b \\ c &= -2d \end{aligned}$$

So,

$$M = \begin{bmatrix} -b & b \\ -2d & d \end{bmatrix}$$

We need to choose  $b, d$  such that  $\det M \neq 0$ . Now,

$$\det M = -bd + 2bd = bd$$

So, we can choose  $b = d = 1$ . We get

$$M = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix} \in \text{GL}_2(\mathbb{R})$$

for which, by above calculations,

$$A = MBM^{-1}$$

holds, so  $A$  and  $B$  are conjugate.

## Result

They are conjugate.

Let

$$A = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$$

We will try to find  $M \in \mathrm{SL}_2(\mathbb{R})$ ,

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

, such that

$$A = MBM^{-1}$$

holds. Equivalently,

$$AM = MB$$

or, after plugging in  $A, B, M$ ,

$$\begin{bmatrix} a+c & b+d \\ c & d \end{bmatrix} = \begin{bmatrix} a+b & b \\ c+d & d \end{bmatrix}$$

This yields the system of equations

$$\begin{aligned} a+c &= a+b \\ b+d &= b \\ c &= c+d \\ d &= d \end{aligned}$$

Simplifying,

$$\begin{aligned} b &= c \\ d &= 0 \\ d &= 0 \\ d &= d \end{aligned}$$

So,  $d = 0$  and  $b = c$ . So,

$$M = \begin{bmatrix} a & b \\ b & 0 \end{bmatrix}$$

We will try to find  $a, b$  such that  $\det M = 1$ . This yields

$$\det M = 1 \iff a - b^2 = 1 \iff \boxed{a = b^2 + 1}$$

So, picking  $b = 1$  and  $a = 2$ , we get

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$$

which is in  $\mathrm{SL}_2(\mathbb{R})$  and

$$A = MBM^{-1}$$

by above calculations. Thus,  $A$  and  $B$  are conjugate in  $\mathrm{SL}_2(\mathbb{R})$ .

Furthermore, since  $\mathrm{SL}_2(\mathbb{R})$  is a subgroup of  $\mathrm{GL}_2(\mathbb{R})$ , we conclude that  $M \in \mathrm{GL}_2(\mathbb{R})$ , so  $A$  and  $B$  are also conjugate in  $\mathrm{GL}_2(\mathbb{R})$ .

## Result

They are conjugate in both groups.



### Subset?

Let  $x = ghg^{-1} \in gHg^{-1}$ . Then, since  $H$  is a subgroup of  $G$ ,  $h \in G$ . Moreover, since  $G$  is a group, we now conclude that  $ghg^{-1} \in G$ . Thus,  $gHg^{-1} \subseteq G$  is true.

### Closure?

Let  $x, y \in gHg^{-1}$ ,  $x = ghg^{-1}$ ,  $y = gkg^{-1}$ , for some  $h, k \in H$ . Now,

$$xy = ghg^{-1}gkg^{-1} = g(hk)g^{-1}$$

Since  $H$  is a subgroup of  $G$ , we conclude that  $hk \in H$ . Thus,  $xy \in gHg^{-1}$ .

### Inverse?

Let  $x \in gHg^{-1}$ ,  $x = ghg^{-1}$  for some  $h \in H$ . Since  $H$  is a subgroup,  $h^{-1} \in H$ , so  $y = gh^{-1}g^{-1} \in gHg^{-1}$ . Furthermore,

$$xy = ghg^{-1}gh^{-1}g^{-1} = gh(h^{-1})g^{-1} = gg^{-1} = 1$$

and

$$yx = gh^{-1}g^{-1}ghg^{-1} = gh^{-1}hg^{-1} = gg^{-1} = 1$$

Thus,  $y = x^{-1}$ , so  $x^{-1} \in gHg^{-1}$ .

## Result

Check the properties from the definition of a subgroup.

## 8. a

Denote by  $f$  the described mapping.

### homomorphism?

Let  $A, B \in \text{GL}_n(\mathbb{R})$ . Then

$$f(AB) = ((AB)^t)^{-1} = (B^t A^t)^{-1} = (A^t)^{-1} (B^t)^{-1} = f(A)f(B)$$

Here we used that

$$(AB)^t = B^t A^t$$

and

$$(AB)^{-1} = B^{-1}A^{-1}$$

Thus,  $f$  is a homomorphism.

### injective?

Suppose that  $f(A) = f(B)$ . Then

$$(A^t)^{-1} = (B^t)^{-1}$$

Multiply by  $A^t B^t$  to get

$$A^t = B^t$$

Transpose both sides and use that  $(M^t)^t = M$ :

$$(A^t)^t = (B^t)^t \implies A = B$$

Thus,  $f$  is injective.

### $f$ surjective?

Let  $X \in \text{GL}_2(\mathbb{R})$ . Then  $(X^{-1})^t \in \text{GL}_2(\mathbb{R})$ , and

$$f((X^{-1})^t) = (((X^{-1})^t)^t)^{-1} = (X^{-1})^{-1} = X$$

Thus,  $f$  is also surjective.

### Conclusion.

Since  $f$  is a bijective homomorphism, it is an isomorphism. Also, its domain and codomain is the same, so it is an automorphism.

## Result

2 of 2

Prove that this mapping is a bijective homomorphism.

9. a

Notice that  $g \mapsto g$  would be a bijection. However, this would not be a homomorphism because in  $G^\circ$  the order of "multiplication" is reversed! We can correct this by defining

$$f : G \rightarrow G^\circ, \quad f(g) = g^{-1}$$

Denote by  $*$  the law of composition of  $G^\circ$ .

### $f$ homomorphism?

Let  $g, h \in G$ . Then

$$f(gh) = (gh)^{-1} = h^{-1}g^{-1} = g^{-1} * h^{-1} = f(g) * f(h)$$

Thus,  $f$  is a homomorphism.

### $f$ injective?

Suppose that  $f(g) = f(h)$ . Then  $g^{-1} = h^{-1}$ . Since the inverse is unique,  $g = h$ . Thus,  $f$  is injective.

### $f$ surjective?

Let  $g \in G^\circ$ . Then  $f(g^{-1}) = (g^{-1})^{-1} = g$ . Thus,  $f$  is surjective.

### Conclusion.

So,  $f$  is a bijective homomorphism, so it is also an isomorphism, which means that  $G \approx G^\circ$ , as required.

## Result

Hint: consider  $f : G \rightarrow G^\circ, f(g) = g^{-1}$ .

10. a

**(a)**

Let

$$G = \langle a \rangle = \{1, a, \dots, a^9\}$$

Let  $\varphi(a) = a^k$ . We will prove that

$$\text{im}\varphi = \langle a^k \rangle$$

First of all, if  $y \in \text{im}\varphi$ , then there exists some  $g \in G$  such that

$$\varphi(g) = y$$

However,  $g = a^l$ , for some  $0 \leq l < 10$ , so

$$y = \varphi(a^l) = \underbrace{\varphi(a \cdots a)}_{l \text{ times}} = \underbrace{\varphi(a) \cdots \varphi(a)}_{l \text{ times}} = \underbrace{a^k \cdots a^k}_{l \text{ times}} = (a^k)^l \in \langle a^k \rangle$$

Thus,

$$\text{im}\varphi \subseteq \langle a^k \rangle$$

Now take some  $x \in \langle a^k \rangle$ . Then there exists some integer  $l$  such that  $x = (a^k)^l$ . The previous computations now show that

$$\varphi(a^l) = (a^k)^l = x,$$

so  $x \in \text{im}\varphi$ , and

$$\text{im}\varphi = \langle a^k \rangle$$

Since  $a^k$  generates the entire  $G$  if and only if  $\gcd(k, 10) = 1$  (Proposition 2.4.3.), for  $a^k$  to generate the entire  $G$  (and for  $\varphi$  to be an automorphism!), we must have  $k = 1, 3, 7, 9$ . Thus, there are 4 automorphisms.

**(b)**

First notice that

$$S_3 = \{\text{id}, (12), (123), (123)^2, (12)(123), (123)(12)\}$$

Now let  $\varphi$  be an automorphism. Since it is injective and  $\varphi(\text{id}) = \text{id}$ , we must have  $\varphi((12)) \neq \text{id}$ . Moreover, we must send it to some element of order 2! To prove that, let  $g = \varphi((12))$ , then

$$g^2 = \varphi((12)^2) = \varphi(\text{id}) = \text{id}$$

Similarly, suppose that  $h = \varphi((123))$ . First of all,  $\varphi$  being injective,  $h \neq \text{id}$ . Now suppose that  $h^2 = \text{id}$ . Then

$$h^2 = \varphi((123)^2)$$

Since  $(123)^2 = (132) \neq \text{id}$ , this is a contradiction with the injectivity of  $\varphi$ .

Now all that is left to notice is that

$$S_3 = \{\text{id}, g, h, h^2, gh, hg\}$$

where  $g$  is **any** element of  $S_3$  of order 2 and  $h$  is **any** element of  $S_3$  of order 3. Therefore, if we define

$$\begin{aligned}\varphi(\text{id}) &= \text{id} \\ \varphi((12)) &= g \\ \varphi((123)) &= h \\ \varphi((123)^2) &= h^2 \\ \varphi((12)(123)) &= gh \\ \varphi((123)(12)) &= hg\end{aligned}$$

we get one automorphism.

### Result

(a) Any homomorphism  $\varphi$  with  $\varphi(a) = a^k$ , where  $k = 1, 3, 7, 9$ .

(b) Let  $g$  be of order 2 and  $h$  be of order 3. Then an automorphism is given by

$$\begin{aligned}\varphi(\text{id}) &= \text{id} \\ \varphi((12)) &= g \\ \varphi((123)) &= h \\ \varphi((123)^2) &= h^2 \\ \varphi((12)(123)) &= gh \\ \varphi((123)(12)) &= hg\end{aligned}$$

## 11. a

If  $a = 1$ , then the statement is trivial. For the rest of the proof, assume  $a \neq 1$ .

We must prove that  $a$  commutes with every element  $g$  of  $G$ ; that is,

$$ga = ag$$

Now let  $g \in G$  be arbitrarily taken and denote  $N = \{1, a\}$ . Since  $N$  is a normal subgroup,

$$gag^{-1} \in N$$

So,  $gag^{-1} = 1$  or  $gag^{-1} = a$ . For the first case,

$$gag^{-1} = 1 \Leftrightarrow ga = g \Leftrightarrow a = 1$$

(in the first equality multiply the equation from the right by  $g$ , in the second multiply it from the left by  $g^{-1}$ ).

However, we assumed  $a \neq 1$ , so we must have that  $gag^{-1} = a$ . Multiplying from the right by  $g$  yields

$$ag = ga$$

as required.

### Result

2 of 2

Prove that  $ga = ag$  for every  $g \in G$  (hint:  $gag^{-1} \in \{1, a\}$ ).

## Section 7

1. a

Let  $G$  be a group and  $a \sim b$  if  $a = bgb^{-1}$  for some  $g \in G$ . The relation is equivalence relation.

1.  $a \sim a$  since  $a = eae^{-1}$  for identity element  $e \in G$ . The relation is reflexive.
2. If  $a \sim b$  then  $a = bgb^{-1}$  for some  $g$ .

$$a = bgb^{-1} \implies ag = gb \implies g^{-1}ag = b$$

Now take  $g' = g^{-1}$  then we get

$$b = g'a(g')^{-1}$$

which shows  $b = gag^{-1}$ . i.e. the relation is symmetric.

3. Now suppose  $a \sim b$  and  $b \sim c$  then there exists  $g, h \in G$  such that  $a = bgb^{-1}$  and  $b = hch^{-1}$ . This gives

$$a = g(hch^{-1})g^{-1} \implies a = (gh)c(gh)^{-1}$$

i.e.  $a \sim c$ . This shows that relation is transitive.

Therefore the relation is equivalence relation.

### Result

2 of 2

Show that relation is reflexive, symmetric and transitive. For reflexive  $a \sim a$  which holds trivially. For symmetric, take  $g' = g^{-1}$ . For transitive, assume  $a \sim b$  and  $b \sim c$ . Show that  $a \sim c$  by substituting  $b$  in  $a = bgb^{-1}$ .

2. a

Given  $R$  is subset of  $S \times S$ . The relation  $R$  is equivalence relation if it is reflexive, symmetric and transitive.

1. For reflexive,  $a \sim a \implies (a, a) \in R$ .
2. For symmetric,  $a \sim b \implies b \sim a$  so  $(a, b) \in R$  should imply  $(b, a) \in R$ .
3. For transitive  $a \sim b, b \sim c$  implies  $a \sim c$  so  $(a, b), (b, c) \in R$  should imply  $(a, c) \in R$ .

### Result

2 of 2

(i)  $(a, a) \in R$ , (ii)  $(a, b) \in R \implies (b, a) \in R$  and (iii)  $(a, b), (b, c) \in R$  implies  $(a, c) \in R$ .

3. a



Let  $R$  and  $R'$  be two equivalence relation on  $S$ . The intersection  $R \cap R'$  is equivalence relation. Since both  $R$  and  $R'$  is reflexive,  $R \cap R'$  is also reflexive. If  $(a, b) \in R \cap R'$  then  $(b, a) \in R \cap R'$  as both set contain this element. Similarly  $(a, b), (b, c) \in R \cap R'$  then  $(a, c) \in R, R'$  hence  $(a, c) \in R \cap R'$ . This shows  $R \cap R'$  is transitive.

However  $R \cup R'$  is not equivalence relation. Let  $S = \{1, 2, 3\}$  and  $R = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1)\}$  and  $R' = \{(1, 1), (2, 2), (3, 3), (2, 3), (3, 2)\}$  then  $R \cup R'$  contains  $(1, 2)$  and  $(2, 3)$  however it does not contain  $(1, 3)$ .

## Result

2 of 2

$R \cap R'$  is equivalence relation however  $R \cup R'$  is not equivalence relation.

4. a

Geometrically, symmetric means that elements of  $R$  must contain it's mirror image about line  $y = x$ . Mirror image of  $(a, b)$  about line  $y = x$  is  $(b, a)$ . Reflexive means  $R$  must contain all points of line  $y = x$ .

5. a

(a) For  $\{(s, s) : s \in \mathbb{R}\}$ , all axioms are satisfied by this relation since it contains only points of  $y = x$  line which is reflexive. It is naturally symmetric and transitive.

(b) For empty set, the relation is not reflexive. It is symmetric and transitive trivially since there are not elements.

## Step 2

2 of 3

(c) For the locus  $\{xy + 1 = 0\}$ , the relation is not reflexive since it does not intersect with  $y = x$ . The relation is symmetric since  $(x, y)$  lines in the locus implies  $(y, x)$  also lines in the locus. Now, Now,

$$xy + 1 = 0 \implies y = -\frac{1}{x}$$

For  $x \neq 0$ , this point is unique so  $(x, -\frac{1}{x})$  lines in the line. Now for  $-\frac{1}{x}$ , we get

$$-1/(-\frac{1}{x}) = x$$

so  $(-\frac{1}{x}, x)$  is the only other point. Hence relation is transitive.

(d) Given the locus  $\{x^2y - y^2x - x + y = 0\}$ . For any  $(s, s)$ , we see that  $s^2s - s^2s - s + s = 0$  so relation is reflexive. The relation is also symmetric since changing  $y \rightarrow x$ , we get

$$y^2x - x^2y - y + x = -(x^2y - xy^2 - x + y) = 0$$

Now, simplifying equation gives

$$xy(x - y) - (x - y) = (xy - 1)(x - y) = 0$$

Now, if  $x \neq y$  then  $y = \frac{1}{x}$  so  $(x, \frac{1}{x})$  is in the locus. For  $x \neq 0$ , the other coordinate is either  $x$  or  $\frac{1}{x}$  which is unique. Now for  $\frac{1}{x}$ , the other coordinate is either  $\frac{1}{x}$  or  $1/(\frac{1}{x}) = x$ . Hence the relation is transitive.

## Result

(a) All (b) Not reflexive (c) Not reflexive (d) All

6. a

The number of equivalence relation on set correspond to the number of partition of the set. The partition of integer 5 is 52 which is also Bell's number  $B_5$ . Thus there are 52 possible equivalence relations on set  $S$  consisting of five elements.

## Result

2 of 2

52 possible equivalence relations.

## Section 8

1. a

Given  $H = \{(), (1, 2, 3), (1, 3, 2)\}$  is subgroup of alternating group  $A_4 = \{(), (1, 3)(2, 4), (1, 2)(3, 4), (1, 4)(2, 3), (2, 4, 3), (1, 3, 4), (1, 2, 3), (1, 4, 2), (2, 3, 4), (1, 3, 2), (1, 2, 4), (1, 4, 3)\}$ . The right cosets are calculated as

$$\begin{aligned} H() &= \{(), (1, 2, 3), (1, 3, 2)\} \\ H(2, 3, 4) &= \{(2, 3, 4), (1, 3)(2, 4), (1, 4, 2)\} \\ H(2, 4, 3) &= \{(2, 4, 3), (1, 4, 3), (1, 2)(3, 4)\} \\ H(1, 2, 4) &= \{(1, 2, 4), (1, 4)(2, 3), (1, 3, 4)\} \end{aligned}$$

The left cosets are calculated as

$$\begin{aligned} ()H &= \{(), (1, 2, 3), (1, 3, 2)\} \\ (2, 3, 4)H &= \{(2, 3, 4), (1, 2)(3, 4), (1, 3, 4)\} \\ (2, 4, 3)H &= \{(2, 4, 3), (1, 2, 4), (1, 3)(2, 4)\} \\ (1, 4, 2)H &= \{(1, 4, 2), (1, 4, 3), (1, 4)(2, 3)\} \end{aligned}$$

## Result

2 of 2

The right cosets of  $H$  are  $\{(), (1, 2, 3), (1, 3, 2)\}, \{(2, 3, 4), (1, 3)(2, 4), (1, 4, 2)\}, \{(2, 4, 3), (1, 4, 3), (1, 2)(3, 4)\}, \{(1, 2, 4), (1, 4)(2, 3), (1, 3, 4)\}$  and  $\{(), (1, 2, 3), (1, 3, 2)\}, \{(2, 3, 4), (1, 2)(3, 4), (1, 3, 4)\}, \{(2, 4, 3), (1, 2, 4), (1, 3)(2, 4)\}, \{(1, 4, 2), (1, 4, 3), (1, 4)(2, 3)\}$  respectively.

2. a



Let  $\mathbb{R}^m$  is additive group of vectors and  $W$  be the set of solution of a system of homogeneous linear equations  $AX = 0$ . Then  $W$  is subgroup of  $\mathbb{R}^m$ . For inhomogeneous equation  $AX = B$ , let  $V$  represent it's solution. If the system  $AX = B$  is inconsistent, it has no solution hence  $V$  is empty. Otherwise, let  $X'$  be a solution of  $AX = B$ . Then for any  $c \in \mathbb{R}$  and  $X'' \in W$

$$A(X' + cX'') = AX' + cAX'' = B$$

Hence solution is of the form  $X' + W$  which is (additive) coset of  $W$ .

## Result

2 of 2

If the system is consistent, show that  $X' + W$  is solution of  $AX = B$  for some solution  $X'$ .

## 3. a

**(Cauchy's Theorem)** If  $p$  is a prime number that divides order of group  $G$  then  $G$  contains an element of order  $p$ .

If  $G$  is of order  $p^n$  then there exists element of order  $p$ .

This can also be proved using **induction**. Suppose group has order prime  $p$  then it is cyclic group which has generating element of order  $p$ . Suppose, for all  $k \leq n$ , group of order  $p^k$  has an element of order  $p$ .

For  $|G| = p^{n+1}$ , suppose if  $G$  has an element  $a$  of order  $p^{n+1}$ , then we may take  $a^{p^n}$  which is not identity and has order  $p$ . If  $G$  does not have an element of order  $p^{n+1}$  (since order of subgroup divides order of group) then take any  $e \neq b \in G$  then  $\langle b \rangle$  has an order of  $p^m$  where  $m \leq n$ . By induction, it has element of order  $p$ .

## Result

2 of 2

By Cauchy's theorem there exists an element of order  $p$ .

## 4. a

**(Cauchy's Theorem)** If  $p$  is a prime number that divides order of (finite) group  $G$  then  $G$  contains an element of order  $p$ .

Since 7 and 5 are primes dividing 35, by Cauchy's theorem there exists element of order 7 and 5.

**Another method**: The order of element must divide the order of group. Since  $7 \cdot 5 = 35$ , the order of (non identity) element must be 7 or 5.

Suppose if  $G$  has order 35 and it does not have element of order 7, then all element must be of order 5. Consider the subgroups which is of order 5. There are five elements in each with one common identity element so there must be  $4n + 1$  element of  $G$ . But for any  $n \in \mathbb{Z}_+$ ,  $4n + 1 \neq 35$  thus it contradicts that  $G$  is of order 35.

Similarly if we assume that there are no element of order 5, then we get  $6n + 1 = 35$  for which  $n$  has no solution in  $\mathbb{Z}_+$ .

Therefore  $G$  must have elements of order 5 as well as 7.

## Result

2 of 2

By Cauchy's theorem, there exists elements of order 5 and 7.

5. a

Since  $G$  is finite, the orders of  $x$  and  $y$  divide the order of  $G$ . Let  $n = |G|$ . Then, since the order of  $x$  is 10,

$$10|n,$$

so there exists a positive integer  $k$  such that

$$n = 10k$$

Moreover, since 6 is the order of  $y$ , and  $6|n$ , there exists a positive integer  $l$  such that

$$n = 6l$$

So,  $6l = 10k$ , or, after dividing by 2,  $3l = 5k$ . Therefore,  $3|5k$ . Since 3 is prime,  $3|5$  or  $3|k$ . However, clearly  $3 \nmid 5$ , so  $3|k$ , which means that there exists some positive integer  $m$  such that

$$k = 3m \Rightarrow n = 30m$$

(since  $n = 10k$ ). This also means that

$$|G| = 30m$$

## Result

2 of 2

$$|G| = 30m, \text{ for some positive integer } m.$$

6. a

By Corollary 2.8.13,

- $|G| = |\ker \varphi| \cdot |\operatorname{im} \varphi|$
- $|\ker \varphi| \mid |G|$
- $|\operatorname{im} \varphi| \mid |G|, |\operatorname{im} \varphi| \mid |G'|$

Thus, if  $|\ker \varphi| = n$ ,  $|\operatorname{im} \varphi| = m$ ,

- $nm = 18$
- $n \mid 18$
- $m \mid 18, m \mid 15$

We will start with the last condition. The divisors of 18 are: 1, 2, 3, 6, 9, 18. The divisors of 15 are: 1, 3, 5, 15. So,  $m = 1$  or  $m = 3$  (since  $m$  divides both 15 and 18, and 1 and 3 are the only positive integers with this property). However,  $|\operatorname{im} \varphi| = m = 1$  would mean that  $\varphi$  is a trivial homomorphism (since the identity is always in the image of a homomorphism). Therefore,

$$|\operatorname{im} \varphi| = m = 3$$

Now we just use the first condition:

$$nm = 18 \implies 3n = 18 \implies |\ker \varphi| = n = 6$$

## Result

$$|\ker \varphi| = 6$$

7. a

Let  $H = \langle x, y \rangle$  be the subgroup of  $G$  generated by  $x$  and  $y$ . To prove that  $H = G$ , it is sufficient to prove that  $|H| = 22$  (because then  $H \subseteq G$  and  $|H| = |G|$  yield  $H = G$ ).

Since  $H$  is a subgroup of  $G$ , by Lagrange's Theorem,  $|H| \mid |G|$ . So,  $|H| \mid 22$ . Since the divisors of 22 are 1, 2, 11, 22, we conclude that  $|H| = 1$ ,  $|H| = 2$ ,  $|H| = 11$ , or  $|H| = 22$ . We now prove that all cases except for the last lead to contradiction.

If  $|H| = 1$ , then  $H$  is trivial; that is, only 1 is in  $H$ . However this is not possible, since  $x \in H$  and  $x \neq 1$ .

If  $|H| = 2$ , then, because  $1, x \in H$ ,  $H = \{1, x\}$ . Since  $H$  is a subgroup of  $G$  and  $x \in H$ ,  $x^2 = x \cdot x$  is in  $H$ . So, we must have  $x^2 = 1$  (because the only other case is  $x^2 = x$ , which yields  $x = 1$ , a contradiction). Moreover,  $y \in H$  by definition of  $H$ . This means that  $y = x$  or  $y = 1 = x^2$ . However, by the assumption of the exercise,  $y$  is not a power of  $x$ . Contradiction.

If  $|H| = 11$ , then  $H = \langle x \rangle$ . This is because  $H$  is a group itself, and  $x \in H$ , so  $|x| \mid |H|$ . So,  $|x| = 1$  or  $|x| = 11$ . Since  $x \neq 1$ , we conclude that  $|x| = 11$ . This means that  $|\langle x \rangle| = 11$ . Thus, truly  $H = \langle x \rangle$ . As before, now  $y = x^k$ , for some  $k = 1, 2, \dots, 11$  (because  $x^{11} = 1$ ), which is a contradiction.

Therefore,  $|H| = 22$  is the only possible outcome, which prove that  $H = G$ .

## Result

2 of 2

Let  $H = \langle x, y \rangle$ . Since  $H$  is a subgroup of  $G$ , and  $|G| = 22$ , by Lagrange's Theorem  $|H| = 1, 2, 11$ , or 22. Show that the first three cases lead to a contradiction, thus  $|H| = 22$ . Conclude that  $H = G$ .

## 8. a

Let  $x \in G$ ,  $x \neq 1$ . Then  $|x| \mid |G|$ , so  $|x| \mid 25$ . This means that  $|x| = 1, 5$ , or 25. Since  $x \neq 1$ ,  $|x| \neq 1$ .

If  $|x| = 5$ , then by definition  $\langle x \rangle$  is of order 5, and it is a subgroup of  $G$ , so we are done.

If  $|x| = 25$ , then consider  $H = \langle x^5 \rangle$ . Now we see that

$$H = \{1, x^5, x^{10}, x^{15}, x^{20}\},$$

since  $x^{25} = 1$ . Thus,  $H$  is a subgroup of  $G$  of order 5.

Now suppose that  $G$  contains only one subgroup of order 5; denote it by  $H$ . Suppose that  $G$  is not cyclic.

Let  $x \in G$ ,  $x \neq 1$ . By Lagrange's Theorem,  $|x| \mid 25$ . Thus,  $|x| = 1, 5$ , or 25. Since  $x \neq 1$ ,  $|x| \neq 1$ . Moreover,  $|x| = 25$  would mean that  $\langle x \rangle$  was of order 25, so  $G = \langle x \rangle$ , which is a contradiction since we assumed that  $G$  was not cyclic.

Thus,  $|x| = 5$ , so  $H = \langle x \rangle$ . Now notice that  $G \setminus H \neq \emptyset$ ! Thus, there exists some  $y \in G \setminus H$ ; more precisely,  $y \notin \langle x \rangle$ , so  $y \neq 1$ ,  $y \neq x^k$ ,  $k = 1, 2, 3, 4$ . On the other hand, as with  $x$ , we conclude that  $|y| = 5$ . Thus,  $\langle y \rangle$  is another subgroup of  $G$  of order 5. This is a contradiction with the assumption that  $G$  contains only one subgroup of order 5, so our assumption that  $G$  is not cyclic was false.

To conclude,  $G$  truly must be cyclic.

## Result

2 of 2

For the first part, pick any  $x \in G$ ,  $x \neq 1$ . What are the possible orders of  $x$ ? For each case, define the required subgroup.

For the second part, show that if  $G$  is not cyclic, we can find at least two different subgroup of order 5.

## 9. a

Suppose that  $\varphi$  is an automorphism. First of all, it must be a homomorphism, so

$$\varphi(xy) = \varphi(x)\varphi(y),$$

for all  $x, y \in G$ . This means that

$$(xy)^2 = x^2y^2 \Rightarrow xyxy = xxyy$$

Multiply the above equality by  $x^{-1}$  from the left and by  $y^{-1}$  from the right to get

$$xy = yx,$$

for all  $x, y \in G$ . Thus,  $G$  must be commutative.

Moreover,  $\varphi$  must be bijective. For it to be bijective, it is necessary for it to be injective. Now suppose that  $x \in G$  is of even order,  $|x| = 2k$ . Then

$$\varphi(x^k) = x^{2k} = 1$$

Since  $x^k \neq 1$  (if  $x^k = 1$ , then  $|x| = k$ , not  $|x| = 2k$ , since  $k < 2k$ ), and  $\varphi(1) = 1$ , we conclude that  $\varphi$  is not injective.

Therefore, in  $G$  there exists no element of even order.

Now we prove the converse; if  $G$  is commutative such that there exists no element of even order in  $G$ , then  $\varphi$  is an automorphism.

First of all,

$$\varphi(xy) = (xy)^2 = xyxy = xxyy = x^2y^2 = \varphi(x)\varphi(y)$$

for all  $x, y \in G$ , so  $\varphi$  is a homomorphism.

Now let  $x \in G, x \neq 1$ . Since  $|x|$  is not even, specially  $|x| \neq 2$ . This also means that

$$\varphi(x) = x^2 \neq 1, \quad \text{for all } x \neq 1 \quad (1)$$

Now suppose that  $\varphi(x) = \varphi(y)$ , for some  $x, y \in G$ . Then

$$x^2 = y^2 \implies x^2y^{-2} = 1$$

Moreover,

$$\varphi(xy^{-1}) = (xy^{-1})(xy^{-1}) = x^2y^{-2},$$

so

$$\varphi(xy^{-1}) = 1$$

Now, from (1) we can conclude that

$$xy^{-1} = 1 \implies x = y$$

Therefore,  $\varphi$  is injective. Since it is a function on a finite set to itself, it must also be surjective; hence, it is bijective.

Finally,  $\varphi$  is trully automorphism with the given conditions.

## Result

$\varphi$  is an automorphism if and only if  $G$  is commutative and  $G$  contains no elements of even order.



Let  $H$  be a subgroup of  $G$  of index 2. Let  $g \in G$ . We will show that

$$gH = Hg$$

There are two cases. If  $g \in H$ , then

$$gH = \{gh \mid h \in H\} = H$$

$$Hg = \{hg \mid h \in H\} = H$$

since  $H$  is a subgroup of  $G$ . Thus,

$$gH = Hg$$

Now suppose that  $g \notin H$ . Then  $gH \neq H$ , so  $H$  and  $gH$  are two left cosets of  $G$ . Since  $[G : H] = 2$ , these are the only left cosets, so

$$G = H \cup gH$$

since left cosets form a partition of  $G$ .

Now,  $Hg = \{hg \mid h \in H\}$ . Thus, we take  $h \in H$ , and show that  $hg \in gH$ . Suppose that  $hg \notin gH$ . Then  $hg \in H$  (since  $hg \in G$  and  $G = H \cup gH$ !), so there exists some  $h' \in H$  such that

$$hg = h'$$

This also means that

$$g = h^{-1}h' \in H,$$

which is a contradiction with  $g \notin H$ . Thus,  $hg \in gH$ , and  $Hg \subseteq gH$ , since  $hg \in gH$  holds for all  $h \in H$ .

The other inclusion is obtained similarly, by noticing that

$$G = H \cup gH$$

Thus,  $gH = Hg$ .

This means that for every  $g \in G$  we have

$$gH = Hg,$$

and by using the Proposition 2.8.17 (iii) we can finally conclude that  $H$  is a normal subgroup.

Now let  $G = S_3$ ,  $H = \{1, (12)\} = \langle (12) \rangle$  (the last equality holds because  $(12)$  is clearly of order 2). Thus,  $H$  is a subgroup of  $G$  of order 2. Now,

$$[G : H] = |G|/|H| = 3$$

So,  $H$  is also of index 3.

We will prove that  $H$  is not a normal subgroup of  $G$ . To do that, notice that

$$(23)(12)(23)^{-1} = (23)(12)(23) = (13) \notin H$$

Thus,

$$(23)H(23)^{-1} \not\subseteq H,$$

so  $H$  is not a normal subgroup of  $G$ .

## Result

3 of 3

Hint: Proposition 2.8.17 (iii).

When  $[G : H] = 3$ , then  $H$  does not have to be normal. Consider  $G = S_3$ ,  $H = \langle (12) \rangle$

## 11. a

We will first see what are the left and right cosets. Let

$$g = \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}$$

,  $x > 0$ . Let

$$h = \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}$$

,  $z > 0$ . Then

$$gh = \begin{bmatrix} xz & y \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} u & y \\ 0 & 1 \end{bmatrix}$$

This means that

$$gH = \left\{ \begin{bmatrix} u & y \\ 0 & 1 \end{bmatrix} \mid u > 0 \right\}$$

Therefore,  $gH$  can be considered a horizontal ray  $\{(x, y) \mid x > 0\}$  in the right half-plane.

On the other hand,

$$hg = \begin{bmatrix} zx & zy \\ 0 & 1 \end{bmatrix},$$

so

$$Hg = \left\{ \begin{bmatrix} zx & zy \\ 0 & 1 \end{bmatrix} \mid z > 0 \right\}$$

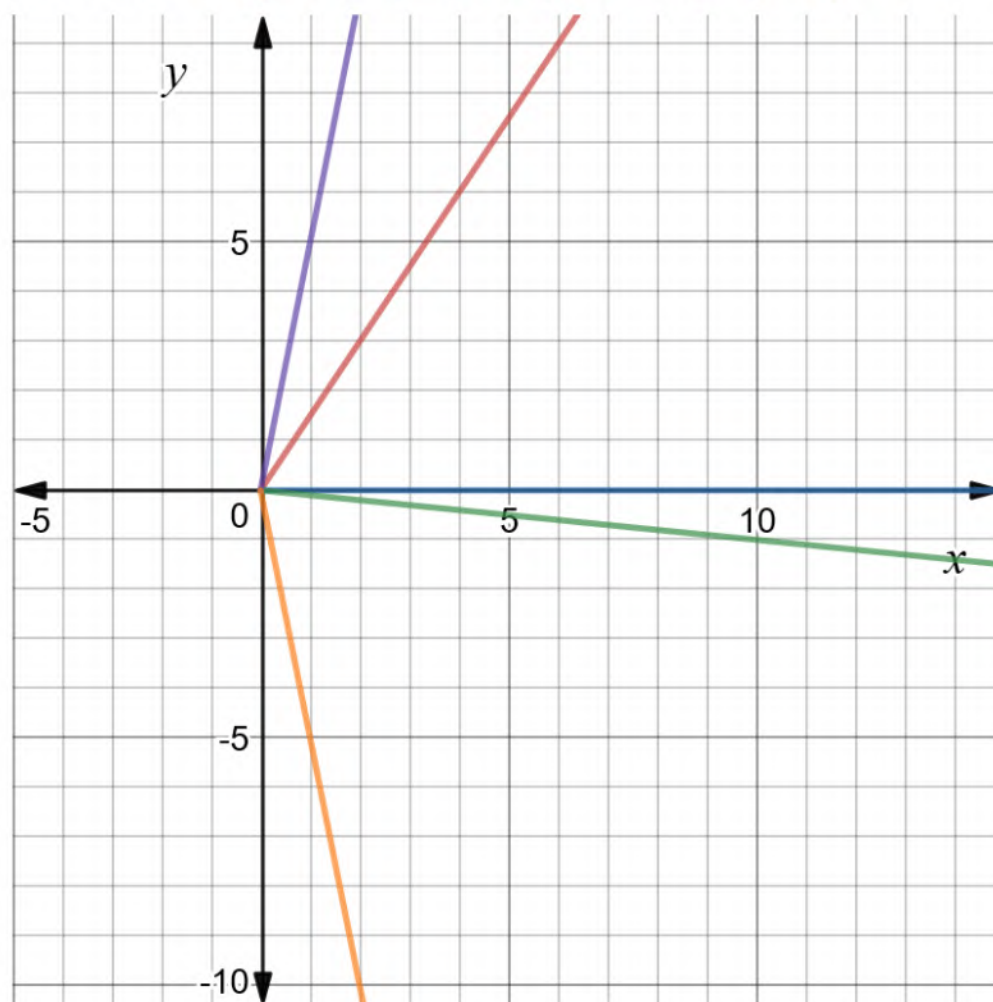
Therefore,  $Hg$  can be considered as a ray  $\{(zx, zy) \mid z > 0\}$ .

Sketches of  $gH$  for  $y = 3$ ,  $y = 0$ ,  $y = -2$ .





Sketches of  $Hg$  for  $x = 2, y = 3, x = 1, y = 0, x = 10, y = -1, x = 1, y = 5, x = 1, y = -5$ :



### Result

Left cosets are horizontal rays in the right half-plane parallel to the  $x$ -axis.

Right cosets are rays  $\{(zx, zy) \mid z > 0\}$  for  $x > 0$ .

12. a

We check the properties from the definition of a subgroup.

Subset?  $S \subseteq G$  is stated in the text of the exercise.

Closure? Let  $x, y \in S$ . We will first show that  $xS = S$ .

For that, notice that  $x \in xS$  since  $1 \in S$ , and

$$xS = \{xs \mid s \in S\} = \{1\} \cup \{xs \mid s \in S \setminus \{1\}\}$$

Thus,  $x \in S \cap xS$ . Since  $aS$  form a partition of  $G$ , two cosets are either equal or disjoint. Since  $xS$  and  $S$  are not disjoint, it follows that  $xS = S$ .

Now  $xy \in xS$ , so  $xy \in S$ , as required.

Inverse? Let  $x \in S$ . We want to show that  $x^{-1} \in S$ .

Consider the set  $x^{-1}S$ . Since  $x \in S$ ,  $x^{-1}x \in x^{-1}S$ , so  $1 \in x^{-1}S$ . Moreover,  $1 \in S$ . Thus,  $1 \in x^{-1}S \cap S$ , so  $x^{-1}S$  and  $S$  are not disjoint. Since  $aS$  form a partition of  $G$ , we must have that  $x^{-1}S = S$ .

Now, since  $1 \in S$ ,  $x^{-1} = x^{-1} \cdot 1 \in x^{-1}S$ , so  $x^{-1} \in S$ , as required.

## Result

2 of 2

Check the properties from the definition of a subgroup.

13. a

(a)

Operation +.

We will show that

$$[\text{Pos}] + [\text{Neg}]$$

is not contained in any single subset of this partition. To show this, note that

$$2 \in [\text{Pos}], \quad -3 \in [\text{Neg}], \quad \text{and } 2 + (-3) = -1 \in [\text{Neg}]$$

$$2 \in [\text{Pos}], \quad -1 \in [\text{Neg}], \quad \text{and } 2 + (-1) = 1 \in [\text{Pos}]$$

So,  $[\text{Pos}] + [\text{Neg}]$  contains elements of both  $[\text{Pos}]$  and  $[\text{Neg}]$ , so it cannot be contained in only one set of this partition.

First of all, clearly

$$[\{0\}] \times [\text{Pos}] = [\{0\}]$$

$$[\{0\}] \times [\{0\}] = [\{0\}]$$

$$[\{0\}] \times [\text{Neg}] = [\{0\}]$$

$$[\text{Pos}] \times [\{0\}] = [\{0\}]$$

$$[\text{Neg}] \times [\{0\}] = [\{0\}]$$

Now,

$$[\text{Pos}] \times [\text{Pos}] = \{ab \mid a, b \in \text{Pos}\} = \{ab \mid a, b \text{ are positive integers}\} = [\text{Pos}],$$

since the product of positive integers is a positive integer.

Similarly,

$$[\text{Neg}] \times [\text{Neg}] = \{ab \mid a, b \in \text{Neg}\} = \{ab \mid a, b \text{ are negative integers}\} = [\text{Pos}],$$

since the product of two negative integers is a positive integer.

Furthermore,

$$[\text{Pos}] \times [\text{Neg}] = \{ab \mid a \in \text{Pos}, b \in \text{Neg}\} = [\text{Neg}],$$

since a product of a positive and a negative integer is a negative integer.

Since multiplication on  $\mathbb{Z}$  is commutative,

$$[\text{Neg}] \times [\text{Pos}] = [\text{Pos}] \times [\text{Neg}] = [\text{Neg}]$$

So, all combinations of products of subsets that form this partition are contained in some subset of this partition (moreover, each product is equal to some subset of partition!), so  $\times$  is compatible with this partition.

## Result

2 of 2

$+$  is not compatible, while  $\times$  is.

## Section 9

1. a

We will prove that 2 has a multiplicative inverse if and only if  $n$  is odd.

Suppose that  $n$  is odd. Then  $n = 2k + 1$ , for some positive integer  $k$ . Now,

$$\overline{2} \cdot \overline{k+1} = \overline{2(k+1)} = \overline{2k+2} = \overline{n+1} = \overline{1},$$

by the definition of multiplication in  $\mathbb{Z}/\mathbb{Z}n$ . Thus,  $\overline{k+1}$  is a multiplicative inverse of  $\overline{2}$ .

Now suppose that  $n$  is even. That is,  $n = 2k$ , for some positive integer  $k$ . Now,

$$\overline{2} \cdot \overline{k} = \overline{2k} = \overline{n} = \overline{0}$$

Suppose that  $\overline{2}$  has a multiplicative inverse  $\overline{a}$ . Then

$$\overline{2} \cdot \overline{k} = \overline{0} \implies \underbrace{\overline{a} \cdot \overline{2}}_{\overline{1}} \cdot \overline{k} = \overline{a} \cdot \overline{0} \implies \overline{1} \cdot \overline{k} = \overline{0} \implies \overline{k} = \overline{0}$$

However, clearly  $k \not\equiv 0 \pmod{n}$ , so  $\overline{k} \neq \overline{0}$ , which is a contradiction. Thus,  $\overline{2}$  does not have a multiplicative inverse.

## Result

2 c

It has an inverse if and only if  $n$  is odd.

## 2. a

### Modulo 4

First of all, there are four possibilities on  $a$ :

$$a \equiv 0 \text{ modulo } 4$$

$$a \equiv 1 \text{ modulo } 4$$

$$a \equiv 2 \text{ modulo } 4$$

$$a \equiv 3 \text{ modulo } 4$$

Now we consider all cases independently, and use Lemma 2.9.6 implicitly.

If  $a \equiv 0 \text{ modulo } 4$ , then

$$a^2 \equiv 0^2 \text{ modulo } 4 \implies a^2 \equiv 0 \text{ modulo } 4$$

If  $a \equiv 1 \text{ modulo } 4$ , then

$$a^2 \equiv 1^2 \text{ modulo } 4 \implies a^2 \equiv 1 \text{ modulo } 4$$

If  $a \equiv 2 \text{ modulo } 4$ , then

$$a^2 \equiv 2^2 \text{ modulo } 4 \implies a^2 \equiv 4 \text{ modulo } 4 \implies a^2 \equiv 0 \text{ modulo } 4$$

since  $4 \equiv 0 \text{ modulo } 4$ .

If  $a \equiv 3 \text{ modulo } 4$ , then

$$a^2 \equiv 3^2 \text{ modulo } 4 \implies a^2 \equiv 9 \text{ modulo } 4 \implies a^2 \equiv 1 \text{ modulo } 4$$

since  $9 \equiv 1 \text{ modulo } 4$ .

Thus,  $a^2$  is congruent to either 0 or 1 modulo 4.

## Modulo 8

First of all, there are eight possibilities on  $a$ :

$$a \equiv 0 \text{ modulo } 8$$

$$a \equiv 1 \text{ modulo } 8$$

$$a \equiv 2 \text{ modulo } 8$$

$$a \equiv 3 \text{ modulo } 8$$

$$a \equiv 4 \text{ modulo } 8$$

$$a \equiv 5 \text{ modulo } 8$$

$$a \equiv 6 \text{ modulo } 8$$

$$a \equiv 7 \text{ modulo } 8$$

Now we consider all cases independently, and use Lemma 2.9.6 implicitly.

If  $a \equiv 0 \text{ modulo } 8$ , then

$$a^2 \equiv 0^2 \text{ modulo } 8 \implies a^2 \equiv 0 \text{ modulo } 8$$

If  $a \equiv 1 \text{ modulo } 8$ , then

$$a^2 \equiv 1^2 \text{ modulo } 8 \implies a^2 \equiv 1 \text{ modulo } 8$$

If  $a \equiv 2 \text{ modulo } 8$ , then

$$a^2 \equiv 2^2 \text{ modulo } 8 \implies a^2 \equiv 4 \text{ modulo } 8$$

If  $a \equiv 3 \text{ modulo } 8$ , then

$$a^2 \equiv 3^2 \text{ modulo } 8 \implies a^2 \equiv 9 \text{ modulo } 8 \implies a^2 \equiv 1 \text{ modulo } 8$$

since  $9 \equiv 1 \text{ modulo } 8$ .

If  $a \equiv 4 \text{ modulo } 8$ , then

$$a^2 \equiv 4^2 \text{ modulo } 8 \implies a^2 \equiv 16 \text{ modulo } 8 \implies a^2 \equiv 0 \text{ modulo } 8$$

since  $16 \equiv 0 \text{ modulo } 8$ .

If  $a \equiv 5 \text{ modulo } 8$ , then

$$a^2 \equiv 5^2 \text{ modulo } 8 \implies a^2 \equiv 25 \text{ modulo } 8 \implies a^2 \equiv 1 \text{ modulo } 8$$

since  $25 \equiv 1 \text{ modulo } 8$ .

If  $a \equiv 6 \text{ modulo } 8$ , then

$$a^2 \equiv 6^2 \text{ modulo } 8 \implies a^2 \equiv 36 \text{ modulo } 8 \implies a^2 \equiv 4 \text{ modulo } 8$$

since  $36 \equiv 4 \text{ modulo } 8$ .

If  $a \equiv 7 \text{ modulo } 8$ , then

$$a^2 \equiv 7^2 \text{ modulo } 8 \implies a^2 \equiv 49 \text{ modulo } 8 \implies a^2 \equiv 1 \text{ modulo } 8$$

since  $49 \equiv 1 \text{ modulo } 8$ .

Thus,  $a^2$  is congruent to either 0, 1, or 4 modulo 8.

## Result

$a^2$  is congruent to either 0 or 1 modulo 4.

$a^2$  is congruent to either 0, 1, or 4 modulo 8.

### 3. a

Let  $a$  be of the form

$$a = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \dots + a_1 \cdot 10 + a_0$$

Then  $a_n, a_{n-1}, \dots, a_1, a_0$  are digits of  $a$ . Now we will first prove that  $10^n \equiv 1$  modulo 9 for all nonnegative integers  $n$ .

[Prove  \$10^n \equiv 1\$  modulo 9](#)

We will prove this by induction.

The base case is when  $n = 0$ . However, now  $10^0 = 1$ , and  $1 \equiv 1$  modulo 9 is trivial.

Now suppose that the statement holds when  $n = k$ , where  $k$  is some nonnegative integer.

Let  $n = k + 1$ . Then  $10^{k+1} = 10^k \cdot 10$ . Also,  $10^k \equiv 1$  modulo 9 by the Induction Hypothesis, while  $10 \equiv 1$  modulo 9 is seen by definition. Thus, using Lemma 2.9.6,

$$10^{k+1} \equiv 1^2 \equiv 1 \text{ modulo } 9$$

Therefore, the statement holds for  $n = k + 1$ .

Now, by the Principle of Mathematical Induction, we conclude that the statement indeed holds for all nonnegative integers  $n$ .

[Complete the exercise.](#)

Since  $10^n \equiv 1$  modulo 9, then

$$a_k 10^k \equiv a_k \cdot 1 \equiv a_k \text{ modulo } 9$$

for  $k = 0, 1, \dots, n$  by Lemma 2.9.6. Moreover, by the same Lemma, we now have

$$a \equiv a_n + a_{n-1} + \dots + a_1 + a_0 \text{ modulo } 9,$$

which completes the proof of the exercise.

## Result

2 of 2

Let  $a = a_n \cdot 10^n + \dots + a_1 \cdot 10 + a_0$ . Then  $a_i$  are digits of  $a$ . Prove that  $10^k \equiv 1$  modulo 9 for all nonnegative integers  $k$ . Now use Lemma 2.9.6 to prove the rest.

### 4. a



### Modulo 9

By definition,

$$2x \equiv 5 \text{ modulo } 9 \quad \text{if and only if} \quad 9 \text{ divides } 2x - 5$$

This means that

$$2x - 5 = 9k$$

for some integer  $k$ . Furthermore,

$$2x = 9k + 5 = 9(k - 1) + 14 = 9l + 14,$$

for some integer  $l$ , or

$$2(x - 7) = 9l,$$

which means that  $2x \equiv 5 \text{ modulo } 9$  if and only if  $9$  divides  $2(x - 7)$ .

However,  $\gcd(2, 9) = 1$ , so there exist some integer  $a, b$  such that

$$2a + 9b = 1$$

Multiply by  $(x - 7)$ :

$$2(x - 7)a + 9(x - 7)b = x - 7 \implies 9[(x - 7)b] = (x - 7)[1 - 2a]$$

Thus,  $9$  divides  $x - 7$ . This means that

$$x \equiv 7 \text{ modulo } 9$$

is the only candidate for the solution. However,

$$2x \equiv 14 \equiv 5 \text{ modulo } 9$$

so it is trully a solution. Thus,

$$x \equiv 7 \text{ modulo } 9$$

### Modulo 6

By definition,

$$2x \equiv 5 \text{ modulo } 6 \quad \text{if and only if} \quad 6 \text{ divides } 2x - 5$$

This means that

$$2x - 5 = 6k$$

for some integer  $k$ .

Now notice that for every  $x \in \mathbb{Z}$ ,  $2x - 5$  is odd. On the other hand,  $6k$  is clearly even. Thus,

$$2x - 5 = 6k$$

cannot hold for any integers  $x, k$ . Thus, this congruence equation does not have a solution!

### Result

Modulo 9:  $x \equiv 7 \text{ modulo } 9$ .

Modulo 6: No solution.

5. a



From  $2x - y \equiv 1$  we get  $y \equiv 2x - 1$ . Thus,

$$4x + 3y \equiv 2 \Leftrightarrow 4x + 3(2x - 1) \equiv 2 \Leftrightarrow 10x \equiv 5 \text{ modulo } n$$

Now suppose that  $n$  is such that the solution exists, and denote one solution by  $x_0$ . Then, by definition,

$$10x_0 \equiv 5 \text{ modulo } n \quad \text{if and only if} \quad n \text{ divides } 10x_0 - 5$$

Thus, there exists some integer  $k$ :

$$nk = 10x_0 - 5 \Leftrightarrow nk - 10x_0 = -5$$

Since  $\gcd(n, 10)$  divides  $n$  and 10, it divide the whole left side, so it must also divide the right side; that is, it must divide 5.

Now we want to prove that the above condition is sufficient; that is, that if  $d = \gcd(n, 10)$  divides 5, then

$$10x \equiv 5 \text{ modulo } n$$

has a solution. First of all, we can define integers  $n', a', b'$  such that

$$n = n'd$$

$$10 = a'd$$

$$5 = b'd$$

Thus,

$$10x \equiv 5 \text{ modulo } n \quad \text{if and only if} \quad a'dx \equiv b'd \text{ modulo } n'd$$

By definition, this is if and only if  $n'd$  divides  $(a'dx - b'd) = (a'x - b')d$ , so if and only if there exists an integer  $k$  such that

$$kn'd = (a'x - b')d \Leftrightarrow kn' = (a'x - b'),$$

if and only if  $n'$  divides  $a'x - b'$ . Finally, this is if and only if

$$a'x \equiv b' \text{ modulo } n'$$

Now notice that  $\gcd(a', n') = 1$ . Truly, if  $c$  is a positive common divisor of  $a', n'$ , then  $cd$  is a common divisor of  $0, n$ . Since  $d$  is the greatest common divisor of  $10, n$ , then  $d = cd$ , which means that  $c = 1$ .

Thus, there exist integer  $l, m$  such that

$$a'l + n'm = 1$$

Multiply by  $b'$  to get

$$a'lb' + n'mb' = b' \implies n'(-mb') = a'(lb') - b'$$

Thus,  $n'$  divides  $a'(lb') - b'$ , so

$$a'(lb') \equiv b' \text{ modulo } n'$$

This means that  $x_0 = lb'$  is one solution of

$$a'x \equiv b' \text{ modulo } n'$$

By the previous discussion, now we know that

$$10x_0 \equiv 5 \text{ modulo } n,$$

so the starting equation has a solution.

Finally, it is now clear that the system of equations has a solution

$$x_0, \quad y_0 \equiv 2x_0 - 1$$

## Result

It has a solution if and only if  $\gcd(10, n)$  divides 5.

## 6. a

We will prove that there exists  $x_1$  such that

$$bx_1 \equiv u \text{ modulo } a$$

Since  $\gcd(a, b) = 1$ , there exist integers  $k, l$  such that

$$ak + bl = 1$$

Now multiply the above equation by  $u$ :

$$auk + bul = u \implies a(uk) = b(ul) - u$$

Therefore,  $a$  divides  $b(ul) - u$ , so

$$b(ul) \equiv u \text{ modulo } a$$

So, we can take  $x_1 = ul$ .

Similarly, there exists  $x_2$  such that

$$ax_2 \equiv v \text{ modulo } b$$

Finally, consider  $x_0 = bx_1 + ax_2$ . Then

$$x_0 \equiv bx_1 + ax_2 \equiv bx_1 \equiv u \text{ modulo } a$$

$$x_0 \equiv bx_1 + ax_2 \equiv ax_2 \equiv v \text{ modulo } b$$

(we used that  $ax_2 \equiv 0 \text{ modulo } a$  and  $bx_1 \equiv 0 \text{ modulo } b$ ).

This proves that the starting system has a solution.

## Result

2 of 2

Prove that there exist  $x_1, x_2$  such that  $bx_1 \equiv u \text{ modulo } a$  and  $ax_2 \equiv v \text{ modulo } b$ . What can you tell about  $x_0 = bx_1 + ax_2$ ?

7. a

We will raise the power of  $A$  and  $B$  until we get the identity matrix  $I$ .

$$A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$A^3 = A^2 \cdot A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

So, the order of  $A$  is 3.

$$B^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$B^3 = B^2 \cdot B = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$$

$$B^4 = B^3 \cdot B = \begin{bmatrix} 2 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$B^5 = B^4 \cdot B = \begin{bmatrix} 2 & 2 \\ 2 & 0 \end{bmatrix}$$

$$B^6 = B^5 \cdot B = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$$

$$B^7 = B^6 \cdot B = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

$$B^8 = B^7 \cdot B = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore, the order of  $B$  is 8.

### Result

The order of  $A$  is 3.

The order of  $B$  is 8.

## Section 10

1. a

A permutation cycle is odd if there are even number of elements in it. Similarly, if it has odd number of elements, then it is even cycle. For example,  $(1\ 2\ 3)$  is even cycle and  $(1\ 2\ 3\ 4)$  is odd cycle. Cycle decomposition of a permutation consists of disjoint cycles. If  $n$  is the number of odd disjoint cycles in cycle decomposition of permutation then the parity of  $n$  gives the parity of given permutation.

### Result

2 of 2

The number of disjoint odd cycles gives the parity of permutation.

2. a

(a)

If  $xH \cap yK = \emptyset$ , we are done. Thus, suppose that  $xH \cap yK \neq \emptyset$ , and let  $g \in xH \cap yK$ .

Since  $g \in xH$ , there exists some  $h \in H$  such that  $g = xh$ . Similarly, since  $g \in yK$ , there exists some  $k \in K$  such that  $g = yk$ . We will now prove that  $xH = gH$  and  $yK = gK$ .

To prove that  $xH = gH$ , first take some  $z \in xH$ . It is of the form  $z = xh'$ , for some  $h' \in H$ . Since  $g = xh$ , we get that  $x = gh^{-1}$ , so

$$z = xh' \implies z = gh^{-1}h' = g(\underbrace{h^{-1}h'}_{\in H})$$

$((h^{-1}h') \in H$  holds since  $H$  is a subgroup of  $G$ ).

Thus,  $z \in gH$ , and  $xH \subseteq gH$ .

Now let  $z' \in gH$ . Then there exists some  $h'' \in H$  such that  $z' = gh''$ . Since  $g = xh$ ,

$$z' = gh'' \implies z' = xhh'' = x(\underbrace{hh''}_{\in H})$$

Thus,  $z' \in xH$ , and  $gH \subseteq xH$ .

Finally, we conclude that  $xH = gH$ . Similarly,  $yK = gK$ , so

$$xH \cap yK = gH \cap gK$$

Our next objective is to show that

$$gH \cap gK = g(H \cap K),$$

which truly is a coset of  $H \cap K$ .

Let  $z \in gH \cap gK$ . Then  $z \in gH$ , so  $z = gh$ , for some  $h \in H$ . Similarly,  $z \in gK$ , so  $z = gk$ , for some  $k \in K$ . So,  $gh = gk$ . Dividing by  $g^{-1}$  from the left yields  $h = k$ . So,  $h \in K$ , which, together with  $h \in H$ , yields  $h \in H \cap K$ . Recall that  $z = gh$  to conclude that  $z \in g(H \cap K)$ . Therefore,  $gH \cap gK \subseteq g(H \cap K)$ .

Let  $z \in g(H \cap K)$ . Then there exists  $v \in H \cap K$  such that  $z = gv$ . Since  $v \in H$ , then  $z \in gH$ . Similarly, since  $v \in K$ , then  $z \in gK$ . Therefore,  $z \in gH \cap gK$ . This means that  $g(H \cap K) \subseteq gH \cap gK$ .

Finally,

$$g(H \cap K) = gH \cap gK$$

To complete the exercise,

$$xH \cap yK = gH \cap gK = g(H \cap K)$$

Thus,  $xH \cap yK$  is a coset of  $H \cap K$ .

(b)

We will show (combinatorically) that

$$[G : H \cap K] \leq [G : H][G : K]$$

On the right side, we have all possible combinations of **ordered pairs** of cosets  $(xH, yK)$ .

Now let  $g(H \cap K)$  be some coset of  $H \cap K$ . In (a) we have proved that  $g(H \cap K) = gH \cap gK$ . Thus, this coset "corresponds" (in a counting sense) to  $(gH, gK)$  on the right side.

Since  $[G : H \cap K]$  is a number of cosets of the subgroup  $H \cap K$ , we conclude that

$$[G : H \cap K] \leq [G : H][G : K]$$

This also means that  $[G : H \cap K]$  is now finite.

### Result

(a) Prove that, if  $xH \cap yK \neq \emptyset$  and  $g \in xH \cap yK$ ,

$$xH \cap yK = gH \cap gK = g(H \cap K)$$

(b) Prove that

$$[G : H \cap K] \leq [G : H][G : K]$$

The last equality stated in the hint for (a) may be useful.

### 3. a

The fact that  $\varphi$  is surjective is clear. Let us find the kernel  $K$ :

$$g \in K \Leftrightarrow \varphi(g) = 1$$

Since  $g \in G$ ,  $g = x^j$  for some  $j \in \{0, 1, \dots, 11\}$ . Thus,

$$1 = \varphi(g) = y^j$$

However, since  $y$  is of order 6, we conclude that 6 divides  $j$ . This means that  $j = 0$  or  $j = 6$ . Thus,

$$K = \{x^0, x^6\} = \{1, x^6\}$$

Since  $G$  is cyclic, all subgroups of  $G$  are also cyclic. Now we will directly find all subgroups of  $G$  (just raise the power of  $x^i$  and use that the order of  $x$  is 12 until you get  $x^i$  again):

$$\begin{aligned}
\langle 1 \rangle &= \{1\} \\
\langle x \rangle &= G \\
\langle x^2 \rangle &= \{x^{2k} \mid k \in \mathbb{Z}\} = \{1, x^2, \dots, x^6, \dots, x^{10}\} \\
\langle x^3 \rangle &= \{x^{3k} \mid k \in \mathbb{Z}\} = \{1, x^3, x^6, x^9\} \\
\langle x^4 \rangle &= \{x^{4k} \mid k \in \mathbb{Z}\} = \{1, x^4, x^8\} \\
\langle x^5 \rangle &= \{x^{5k} \mid k \in \mathbb{Z}\} = G \\
\langle x^6 \rangle &= \{x^{6k} \mid k \in \mathbb{Z}\} = \{1, x^6\} = K \\
\langle x^7 \rangle &= \{x^{7k} \mid k \in \mathbb{Z}\} = G \\
\langle x^8 \rangle &= \{x^{8k} \mid k \in \mathbb{Z}\} = \{1, x^4, x^8\} = \langle x^4 \rangle \\
\langle x^9 \rangle &= \{x^{9k} \mid k \in \mathbb{Z}\} = \{1, x^3, x^6, x^9\} = \langle x^3 \rangle \\
\langle x^{10} \rangle &= \{x^{10k} \mid k \in \mathbb{Z}\} = \{1, x^2, \dots, x^6, \dots, x^{10}\} = \langle x^2 \rangle \\
\langle x^{11} \rangle &= \{x^{11k} \mid k \in \mathbb{Z}\} = G
\end{aligned}$$

Thus, all subgroups that contain  $K$  are  $G$ ,  $K$ ,  $\langle x^2 \rangle$ , and  $\langle x^3 \rangle$ .

Now we want to find the correspondence.

$$\begin{aligned}
\varphi(G) &= H && \text{(since } \varphi \text{ is surjective)} \\
\varphi(K) &= \{1\} && \text{(definition of kernel)} \\
\varphi(\langle x^2 \rangle) &= \{\varphi(1), \varphi(x^2), \dots, \varphi(x^{10})\} = \{1, y^2, \dots, y^{10}\} = \{1, y^2, y^4\} = \langle y^2 \rangle \\
\varphi(\langle x^3 \rangle) &= \{\varphi(1), \varphi(x^3), \varphi(x^6), \varphi(x^9)\} = \{1, y^3, y^6, y^9\} = \{1, y^3\} = \langle y^3 \rangle
\end{aligned}$$

(we used that  $y$  is of order 6). Thus, from the Correspondence Theorem, we now define the correspondence  $f$  with

$$\begin{aligned}
f(G) &= H \\
f(K) &= \{1\} \\
f(\langle x^2 \rangle) &= \langle y^2 \rangle \\
f(\langle x^3 \rangle) &= \langle y^3 \rangle
\end{aligned}$$

## Result

$$\begin{aligned}
f(G) &= H \\
f(K) &= \{1\} \\
f(\langle x^2 \rangle) &= \langle y^2 \rangle \\
f(\langle x^3 \rangle) &= \langle y^3 \rangle
\end{aligned}$$

4. a



### Define a bijection.

Since  $\varphi$  is surjective,  $\varphi(G) = G'$ . Also,  $\varphi(H) = H'$ . So, we must prove that

$$[G : H] = [\varphi(G) : \varphi(H)]$$

which can give us an idea which bijection to define.

We will prove that

$$f : \{\text{cosets of } H\} \rightarrow \{\text{cosets of } H'\},$$

defined by

$$f(gH) = \varphi(g)\varphi(H)$$

is a bijection.

### Well-defined.

We first need to prove that  $f$  is a well-defined function. That is, the two elements of the domain yield the same element of the codomain.

Let  $gH = g'H$ . Then

$$f(gH) = f(g'H) \iff \varphi(g)\varphi(H) = \varphi(g')\varphi(H) \iff \varphi(g'^{-1}g)\varphi(H) = \varphi(H)$$

(in the last equality we multiplied the previous equality by  $\varphi(g'^{-1})$  from the left and used that  $\varphi$  is a homomorphism). However,  $gH = g'H$  means that  $g'^{-1}g \in H$ , so  $\varphi(g'^{-1}g) \in \varphi(H)$ , so the last equality indeed holds. Thus,  $f(gH) = f(g'H)$ .

### Injective.

Now to prove that  $f$  is injective. Let  $gH, g'H$  be such that  $f(gH) = f(g'H)$ . By the above discussion, this is if and only if

$$\varphi(g'^{-1}g)\varphi(H) = \varphi(H)$$

This holds if and only if

$$\varphi(g'^{-1}g) \in \varphi(H),$$

which is if and only if

$$\varphi(g'^{-1}g) = \varphi(h),$$

for some  $h \in H$ . From this,

$$\varphi(g'^{-1}gh^{-1}) = 1$$

(multiply by  $\varphi(h^{-1})$  from the right and use that  $\varphi$  is a homomorphism). However, this means that

$$g'^{-1}gh^{-1} \in K,$$

since  $K$  is a kernel of  $\varphi$ . Moreover,  $K \subseteq H$ , so

$$g'^{-1}gh^{-1} \in H$$

That is, there exists some  $h' \in H$  such that

$$g'^{-1}gh^{-1} = h',$$



or

$$g'^{-1}g = h'h$$

So,

$$g'^{-1}g \in H$$

which means that

$$gH = g'H$$

Thus,  $f$  is injective.

$f$  is surjective.

Let  $\tilde{g}H'$  be some coset of  $H'$  in  $G'$ . Since  $\varphi$  is surjective, there exists some  $g \in G$  such that

$$\varphi(g) = \tilde{g}$$

Moreover, we now have

$$f(gH) = \varphi(g)\varphi(H) = \tilde{g}H',$$

since  $H' = \varphi(H)$ . Thus,  $f$  is also surjective.

Conclusion.

Now we know that  $f$  is a bijective function, which means that its domain and codomain have the same number of elements (possibly infinity). By definition of index, this means precisely that

$$[G : H] = [G' : H'],$$

as required.

## Result

Observe

$$f : \{\text{set of cosets of } H\} \rightarrow \{\text{set of cosets of } H'\},$$

defined by

$$f(gH) = \varphi(g)\varphi(H)$$

## 5. a

The elements of  $S_4$  that leaves the partitions

$$\Pi_1 : \{1, 2\} \cup \{3, 4\}, \Pi_2 : \{1, 3\} \cup \{2, 4\}, \Pi_3 : \{1, 4\} \cup \{2, 3\}$$

unchanged is the kernel of the homomorphism. The elements of  $S_4$  that leaves the partition unchanged are

$$K = V_4 = \{(1), (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$$

which is also known as Klein-4 group, is the kernel of group homomorphism.

Now there are six subgroups of  $S_3$  which are

$$\{(1), (1\ 2)\}, \{(1), (1\ 3)\}, \{(1), (2\ 3)\}, \{(1), (1\ 2\ 3), (1\ 3\ 2)\}$$

and the other two groups being itself and trivial group.

The subgroup of  $S_4$  corresponding to trivial group  $\{(1)\}$  is the kernel  $V_4$ .

Now consider subgroup  $\{(1), (1\ 2)\}$  of  $S_3$ . The subgroup of  $S_4$  that contains  $V_4$  and is mapped to this subgroup permutes the partition  $\Pi_1 \leftrightarrow \Pi_2$  which is done by element of  $(2\ 3)$ . Thus the corresponding subgroup of  $S_4$  is

$$K \cup \{(2\ 3)k : k \in K\}$$

whose elements are

$$\{(1), (2\ 3), (1\ 2)(3\ 4), (1\ 2\ 4\ 3), (1\ 4), (1\ 4)(2\ 3), (1\ 3\ 4\ 2), (1\ 3)(2\ 4)\}$$

This group is Dihedral group. Similarly other corresponding subgroups are those containing  $K$  along with  $(1\ 3)$  and  $K$  along with  $(1\ 4)$ . These are groups of order 8.

The subgroup of  $S_3$  corresponding to  $\{(1), (1\ 2\ 3), (1\ 3\ 2)\}$  is Alternating group  $A_3$  which contains kernel  $K$  and the subgroup corresponding to  $S_3$  is  $S_4$  itself.

## Result

The subgroups are Klein-4 group, Dihedral Groups, Alternating ( $A_4$ ) group and  $S_4$  itself.

# Section 11

1. a

It is given that  $x$  has order  $r$  in  $G$  and  $y$  has order  $s$  in  $G'$ . The order of element  $(x, y)$  is given by  $\text{lcm}(r, s)$  where

$$\text{lcm}(r, s) = \frac{rs}{\text{gcd}(r, s)}$$

Since order of  $(x, y)$  in  $G \times G'$  must be divisible by order of  $r$  and  $s$  in  $G$  and  $G'$  respectively.

## Result

2 of 2

The order of  $(x, y)$  is given by least common multiple of  $r$  and  $s$  since order of  $(x, y)$  must be least positive integer whose order is divisible by order of both  $r$  and  $s$ .

2. a

We take  $x = (1\ 2\ 3)$  and  $y = (1\ 2)$ . Since, we have

$K = \langle x \rangle = \{(1), (1\ 2\ 3), (1\ 3\ 2)\}$  and  $H = \langle y \rangle = \{(1), (1\ 2)\}$ . Given  $f : H \times K \rightarrow S_3$  is a map given by

$$f(h, k) = hk$$

Since  $H \cap K = \langle y \rangle \cap \langle x \rangle = \{(1)\}$ , the map is injective. Since  $(1\ 2\ 3)(1\ 2) \neq (1\ 2)(1\ 2\ 3)$ , the map is not homomorphism. Also  $H = \{(1), (1\ 2)\}$  is not normal subgroup of  $S_3$ , however  $HK = S_3$ . The map is not isomorphism since the map is not homomorphism.

## Result

2 of 2

The map is injective but not isomorphic. Also  $HK = G$  but it is not consequence of the Proposition.

### 3. a

Let  $G = \langle x \rangle$  and  $G' = \langle y \rangle$  be two infinite cyclic groups. Then

$$G \times G' = \{(x^m, y^n) : m, n \in \mathbb{Z}\}$$

We show that this group is not cyclic. Let  $e$  and  $e'$  be the identity elements of  $G$  and  $G'$  respectively. Then  $(x^n, e') = (x, e')^n$  and  $(e, y^m) = (e, y)^m$ . However  $(e, y) \neq (x, e')$  and any other element besides  $(e, y)$  and  $(e, y^{-1})$  cannot generate  $(e, y^n)$  and similarly for  $(x^m, e')$ . Hence  $G \times G'$  cannot be cyclic.

## Result

2 of 2

Show that there are two elements in  $G \times G'$  which can ONLY be generated by two different elements. This shows the group is not cyclic.

### 4. a

(a)

Define  $f : H \times K \rightarrow G$ ,  $f(h, k) = hk$ . We check the conditions of Proposition 2.11.4 (d).

$H \cap K = \{1\}$ ?

Let  $x \in H \cap K$ . Since  $x \in H$ ,  $x = \pm 1$ . Since  $x \in K$ ,  $x > 0$ . Thus,  $x = 1$ , so  $H \cap K = \{1\}$ .

$HK = G$ ?

$HK \subseteq G$  is trivial since  $G$  is a group. Now let  $g \in G = \mathbb{R}^\times$ . If  $g > 0$ , then  $g \in K$ , so  $g = 1 \cdot g \in H \times K$ .

If  $g < 0$ , then  $k = -g \in K$ , so  $g = (-1) \cdot k \in H \times K$ .

Thus,  $G \subseteq HK$ , and  $HK = G$ .

$H$  and  $K$  normal subgroups of  $G$ ?

We will show that  $H$  is a normal subgroup of  $G$ ; that is, by Proposition 2.8.17,

$$ghg^{-1} \in H,$$

for every  $g \in G$ ,  $h \in H$ . This is easy to show since  $G$  is abelian; let  $g \in G$ ,  $h \in H$ , and

$$ghg^{-1} = h(gg^{-1}) = h \cdot 1 = h \in H$$

Thus,  $ghg^{-1} \in H$  for every  $g \in G$ ,  $h \in H$ , and  $H$  is a normal subgroup.

We prove that  $K$  is normal using the same arguments.

Conclusion.

By Proposition 2.11.4 (d), we conclude that  $f$  is an isomorphism, so  $G$  is isomorphic to  $H \times K$ .

(b)

We will prove that  $H$  and  $K$  are abelian, while  $G$  is not.

Let  $h_1, h_2 \in H$ ,

$$h_1 = \begin{bmatrix} a_1 & 0 \\ 0 & d_1 \end{bmatrix}, \quad h_2 = \begin{bmatrix} a_2 & 0 \\ 0 & d_2 \end{bmatrix}$$

Then

$$h_1 h_2 = \begin{bmatrix} a_1 a_2 & 0 \\ 0 & d_1 d_2 \end{bmatrix} = \begin{bmatrix} a_2 a_1 & 0 \\ 0 & d_2 d_1 \end{bmatrix} = h_2 h_1$$

Thus,  $H$  is abelian.

Let  $k_1, k_2 \in K$ ,

$$k_1 = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, \quad k_2 = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$$

Then

$$k_1 k_2 = \begin{bmatrix} 1 & a+b \\ 0 & 1 \end{bmatrix} = k_2 k_1$$

Thus,  $K$  is abelian.

Now let

$$g_1 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

$$g_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

. Then clearly  $g_1, g_2 \in G$ , while

$$g_1 g_2 = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$$

$$g_2 g_1 = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$$

Hence,

$$g_1 g_2 \neq g_2 g_1,$$

so  $G$  is not abelian.

Next, we prove that  $H \times K$  is abelian. Let  $(h_1, k_1), (h_2, k_2) \in H \times K$ . Then

$$(h_1, k_1)(h_2, k_2) = (h_1 h_2, k_1 k_2) \stackrel{H, K \text{ abelian}}{=} (h_2 h_1, k_2 k_1) = (h_2, k_2)(h_1, k_1)$$

Thus,  $H \times K$  is abelian.

Now suppose that  $G$  is isomorphic to  $H \times K$ , and let  $f : H \times K \rightarrow G$  be some isomorphism. Let  $g_1, g_2 \in G$ . Then there exist some  $(h_1, k_1) \in H \times K, (h_2, k_2) \in H \times K$  such that  $g_1 = f(h_1, k_1), g_2 = f(h_2, k_2)$ . Then

$$\begin{aligned} g_1 g_2 &= f(h_1, k_1) f(h_2, k_2) \\ &= f((h_1, k_1)(h_2, k_2)) \\ &= f((h_2, k_2)(h_1, k_1)) \\ &= f(h_2, k_2) f(h_1, k_1) \\ &= g_2 g_1 \end{aligned}$$

Thus, we now get that  $G$  is abelian, which is a contradiction. Thus,  $G$  is **not** isomorphic to  $H \times K$ .

(c)

Define  $f : H \times K \rightarrow G, f(h, k) = hk$ . We check the conditions of Proposition 2.11.4 (d).

$H \cap K = \{1\}$ ?

Let  $x \in H \cap K$ . Since  $x \in H$ ,  $|x| = 1$ , and  $x = \cos y + i \sin y$  for some  $y \in \mathbb{R}$ . Since  $x \in K$ ,  $x > 0$ . This means that  $\sin y = 0$ , so  $y = k\pi$ ,  $k \in \mathbb{Z}$ , and that  $\cos y > 0$ . Since  $\cos(k\pi) = \pm 1$ , we conclude that  $\cos y = 1$  (and that  $k = 2l$ ,  $l \in \mathbb{Z}$ ; we however do not need that here). Thus,  $x = 1$ , so  $H \cap K = \{1\}$ .

$HK = G$ ?

$HK \subseteq G$  is trivial since  $G$  is a group. Now let  $g \in G = \mathbb{C}^\times$ .

Since  $g \neq 0$ ,  $|g| \neq 0$ , so we can define  $h = \frac{g}{|g|}$ . Now,  $|h| = 1$ , so  $h \in H$ . Moreover,  $|g| > 0$ , so if we define  $k = |g|$ , we get  $k \in K$ . All that is left is to conclude that  $g = \frac{g}{|g|} \cdot |g| = hk \in HK$ .

Thus,  $G \subseteq HK$ , and  $HK = G$ .

$H$  and  $K$  normal subgroups of  $G$ ?

We will show that  $H$  is a normal subgroup of  $G$ ; that is, by Proposition 2.8.17,

$$ghg^{-1} \in H,$$

for every  $g \in G, h \in H$ . This is easy to show since  $G$  is abelian; let  $g \in G, h \in H$ , and

$$ghg^{-1} = h(gg^{-1}) = h \cdot 1 = h \in H$$

Thus,  $ghg^{-1} \in H$  for every  $g \in G, h \in H$ , and  $H$  is a normal subgroup.

We prove that  $K$  is normal using the same arguments.

#### Conclusion.

By Proposition 2.11.4 (d), we conclude that  $f$  is an isomorphism, so  $G$  is isomorphic to  $H \times K$ .

### Result

(a) Yes. (Hint: Proposition 2.11.4 (d))

(b) No. (Hint: show that  $H \times K$  is abelian, while  $G$  is not.)

(c) Yes. (Hint: Proposition 2.11.4 (d))

### 5. a

This will be proved using a series of equivalent statements.

We have that  $(g_1, g_2)$  is in the center of  $G_1 \times G_2$  if and only if

$$(g_1, g_2)(h_1, h_2) = (h_1, h_2)(g_1, g_2),$$

for every  $(h_1, h_2) \in G_1 \times G_2$ , which is if and only if

$$(g_1h_1, g_2h_2) = (h_1g_1, h_2g_2)$$

This, in turn, holds if and only if

$$g_1h_1 = h_1g_1 \quad \text{and} \quad g_2h_2 = h_2g_2$$

for every  $h_1 \in G_1, h_2 \in G_2$ . So, the statement holds if and only if  $g_1 \in Z_1, g_2 \in Z_2$ , which is equivalent to

$$(g_1, g_2) \in Z_1 \times Z_2$$

Thus,  $(g_1, g_2)$  is in the center of  $G_1 \times G_2$  if and only if  $(g_1, g_2) \in Z_1 \times Z_2$ , so the two sets must be equal.

### Result

2 of 2

Just use the definition of the center.

### 6. a



Let  $H$  be of order 3,  $K$  be of order 5. Then both are cyclic since orders are prime numbers (let  $x \in H$ , then the order of  $x$  divides 3, so  $x$  is of order 1 or 3; if  $x \neq 1$ , then  $x$  generates  $H$ , hence  $H$  is cyclic); let  $H = \langle x \rangle$ ,  $K = \langle y \rangle$ . We will observe

$$HK = \{x^i y^j \mid i = 0, 1, 2, j = 0, 1, 2, 3, 4\}$$

Since  $H$  is normal,  $HK$  is a subgroup of  $G$  by Proposition 2.11.4 (c). So, all that is left is to prove that  $HK$  has 15 elements.

To do that, suppose that

$$x^i y^j = x^k y^l$$

Then

$$x^{i-k} = y^{l-j}$$

Thus,  $x^{i-k} \in H$  and  $x^{i-k} \in K$  (the right side of the above equality is from  $K$ ). Since  $H$  is of order 3, the order of  $x^{i-k}$  divides 3, so  $x^{i-k}$  is of order 1 or 3. Since  $K$  is of order 5, the order of  $x^{i-k}$  divides 5, so  $x^{i-k}$  is of order 1 or 5. Thus,  $x^{i-k}$  is of order 1, so

$$x^{i-k} = 1 \implies x^i = x^k$$

Similarly, by the same arguments,

$$y^j = y^l$$

Thus,

$$x^i y^j = x^k y^l \Leftrightarrow x^i = x^k, y^j = y^l$$

This means that  $HK$  has 15 elements, since

$$x^i \neq x^k, \quad \text{when } i, k \in \{0, 1, 2\}, i \neq k$$

$$y^j \neq y^l, \quad \text{when } j, l \in \{0, 1, 2, 3, 4\}, j \neq l$$

For example, to prove that  $x^i \neq x^k$ , suppose that  $x^i = x^k$ , and let  $i > k$  without loss of generality. Then  $x^{i-k} = 1$ , so  $i - k > 0$ ,  $i - k < 3$ , which is a contradiction with the fact that 3 is the order of  $x$ .

## Result

2 of 2

Let  $H = \langle x \rangle$  be of order 3,  $K = \langle y \rangle$  be of order 5. Consider  $HK$ .

7. a



Given that  $H$  is subgroup of  $G$  and  $\varphi : G \rightarrow H$  is a homomorphism where  $\varphi$  restricted to  $G$  is identity map. Let  $\{Hg : g \in G\}$  represent set cosets of  $H$  in  $G$ . Then  $\varphi(Hg) = \varphi(H)\varphi(g) = H\varphi(g)$ . Since  $\varphi(g) \in H$ , we get  $\varphi(gH) = H$ . Hence for every coset of  $H$  in  $G$ , there is one element which gets mapped to identity hence there is one to one correspondence between cosets of  $H$  in  $G$  with the kernel  $N$  of the map i.e.  $N$  contains one element from every cosets of  $H$  in  $G$ . So the set of all cosets of  $H$  in  $G$  is given by  $\{Hg : g \in N\}$ . Thus by proposition 2.11.4,  $N \cap H = \{0\}$  so the map  $H \times N \rightarrow G$  is injective. By same proposition, the kernel  $N$  is normal subgroup, the map is homomorphism.

Also the set of all cosets of  $H$  in  $G$  is given by  $\{Hg : g \in N\}$ . Every element of  $G$  is contained in some coset so  $HN = G$ . Thus the map is an isomorphism.

## Result

2 of 2

The mapping is injective.

## 8. a

Let  $\Phi : H \rightarrow G \times G'$  be some homomorphism.

Define  $\pi : G \times G' \rightarrow G$  as  $\pi(g, g') = g$ . Similarly, define  $\pi' : G \times G' \rightarrow G'$  as  $\pi'(g, g') = g'$ . We will prove that  $\pi \circ \Phi : H \rightarrow G$ ,  $\pi' \circ \Phi : H \rightarrow G'$  are homomorphisms.

First, prove that  $\pi$  is a homomorphism. Let  $(g_1, g'_1), (g_2, g'_2) \in G \times G'$ . Then

$$\pi((g_1, g'_1)(g_2, g'_2)) = \pi(g_1g_2, g'_1g'_2) = g_1g_2 = \pi(g_1, g'_1)\pi(g_2, g'_2)$$

Now let  $h, h' \in H$ . Then

$$(\pi \circ \Phi)(hh') = \pi(\Phi(hh')) = \pi(\Phi(h)\Phi(h')) = \pi(\Phi(h))\pi(\Phi(h')) = (\pi \circ \Phi)(h)(\pi \circ \Phi)(h')$$

Thus,  $\pi \circ \Phi$  is a homomorphism.

We prove that  $\pi' \circ \Phi$  is a homomorphism the same way.

Thus, we can now define a function

$$f : \{\text{homomorphisms } \Phi : H \rightarrow G \times G'\} \rightarrow \{\text{pairs of homomorphisms } (\varphi, \varphi')\}$$

by

$$f(\Phi) = (\pi \circ \Phi, \pi' \circ \Phi)$$

(we needed to prove that  $\pi \circ \Phi$  and  $\pi' \circ \Phi$  are homomorphisms to see that  $f$  is well-defined). Now we must prove that  $f$  is bijective. We do that by proving that it is both injective and surjective.

\underline{f}injective?

Let  $\Phi_1, \Phi_2$  be such that  $\Phi_1 \neq \Phi_2$ . Then there exists some  $h \in H$  such that  $\Phi_1(h) \neq \Phi_2(h)$  (two **functions**  $f_1, f_2$  with the same domain and codomain are equal if and only if  $f_1(x) = f_2(x)$  for every  $x$  in the domain).

Let

$$\Phi_1(h) = (g_1, g'_1)$$

$$\Phi_2(h) = (g_2, g'_2)$$

Since  $\Phi_1(h) \neq \Phi_2(h)$ ,  $g_1 \neq g_2$  or  $g'_1 \neq g'_2$ . If  $g_1 \neq g_2$ , then

$$(\pi \circ \Phi_1)(h) = \pi(g_1, g'_1) = g_1 \neq g_2 = \pi(g_2, g'_2) = (\pi \circ \Phi_2)(h)$$

Thus,  $\pi \circ \Phi_1 \neq \pi \circ \Phi_2$ . Similarly, if  $g'_1 \neq g'_2$ , then

$$(\pi' \circ \Phi_1)(h) = \pi(g_1, g'_1) = g'_1 \neq g'_2 = \pi(g_2, g'_2) = (\pi' \circ \Phi_2)(h)$$

Thus,  $\pi' \circ \Phi_1 \neq \pi' \circ \Phi_2$ . Hence, in both cases,

$$(\pi \circ \Phi_1, \pi' \circ \Phi_1) \neq (\pi \circ \Phi_2, \pi' \circ \Phi_2),$$

so  $f$  is injective.

\underline{f}surjective?

Let  $(\varphi, \varphi')$  be some pair of homomorphisms  $\varphi : H \rightarrow G$ ,  $\varphi' : H \rightarrow G'$ . Define a function

$$\Phi : H \rightarrow G \times G', \quad \Phi(h) = (\varphi(h), \varphi'(h))$$

We will prove that  $\Phi$  is a homomorphism.

Let  $h_1, h_2 \in H$ . Then

$$\begin{aligned} \Phi(h_1 h_2) &= (\varphi(h_1 h_2), \varphi'(h_1 h_2)) \\ &= (\varphi(h_1) \varphi(h_2), \varphi'(h_1) \varphi'(h_2)) \\ &= (\varphi(h_1), \varphi'(h_1)) (\varphi(h_2), \varphi'(h_2)) \\ &= \Phi(h_1) \Phi(h_2) \end{aligned}$$

Thus,  $\Phi$  is a homomorphism. Now all that is left to see is that

$$\varphi = \pi \circ \Phi, \quad \varphi' = \pi' \circ \Phi$$

For example,

$$(\pi \circ \Phi)(h) = \pi(\Phi(h)) = \pi(\varphi(h), \varphi'(h)) = \varphi(h)$$

Thus,  $(\pi \circ \Phi)(h) = \varphi(h)$  for every  $h \in H$ , so  $\varphi = \pi \circ \Phi$ . The other equality is proven the same way.

Thus,

$$f(\Phi) = (\varphi, \varphi'),$$

so  $f$  is also surjective.

Conclusion.  $f$  is both injective and surjective, so it is also bijective. Thus, it is the desired bijection.

## Result

3 of 3

Let  $\Phi : H \rightarrow G \times G'$  be some homomorphism.

Define  $\pi : G \times G' \rightarrow G$  as  $\pi(g, g') = g$ . Similarly, define  $\pi' : G \times G' \rightarrow G'$  as  $\pi'(g, g') = g'$ .

Prove that

$$f(\Phi) = (\pi \circ \Phi, \pi' \circ \Phi)$$

is the desired bijection.

9. a

$$\underline{HK \text{ subgroup} \implies HK = KH}$$

Let  $HK$  is a subgroup of  $G$ .

Let  $x \in HK$ . Since  $HK$  is a subgroup,  $x^{-1} \in HK$ , and  $x^{-1} = hk$  for some  $h \in H, k \in K$ . Moreover, now

$$x = (x^{-1})^{-1} = (hk)^{-1} = \underbrace{k^{-1}}_{\in K} \underbrace{h^{-1}}_{\in H},$$

hence  $x \in KH$ , and  $HK \subseteq KH$ .

Now let  $x \in KH$ . Then  $x = k'h'$ , for some  $k' \in K, h' \in H$ . Now,

$$x^{-1} = \underbrace{h'^{-1}}_{\in H} \underbrace{k'^{-1}}_{\in K}$$

Thus,  $x^{-1} \in HK$ . Furthermore, since  $HK$  is a subgroup,  $x = (x^{-1})^{-1} \in HK$ . This proves that  $KH \subseteq HK$ .

Finally, we have  $HK = KH$ , as required.

$$\underline{HK = KH \implies HK \text{ subgroup}}$$

Let  $HK = KH$ . We will prove that  $HK$  is a subgroup by checking the properties from the definition of a subgroup.

Subset?  $HK$  is clearly a subset of  $G$  since  $G$  is a group so it is closed to multiplication.

Closure? Let  $x, y \in HK$ . Then  $x = hk, y = h'k'$ , for some  $h, h' \in H, k, k' \in K$ . Now,

$$xy = hkh'k'$$

Now,  $kh' \in KH = HK$ , so there exist some  $h'' \in H, k'' \in K$  such that  $kh' = h''k''$ . Finally,

$$xy = hkh'k' = hh''k''k' = \underbrace{(hh'')}_{\in H} \underbrace{(k''k')}_{\in K}$$

Thus,  $xy \in HK$ .

Inverse? Let  $x \in HK$ . Then  $x = hk$  for some  $h \in H, k \in K$ . Now,

$$x^{-1} = (hk)^{-1} = k^{-1}h^{-1}$$

Thus,  $x^{-1} \in KH$ . However,  $KH = HK$ , so  $x^{-1} \in HK$ .

Conclusion.  $HK$  is a subgroup of  $G$  since all properties hold.

To prove that  $HK = KH$  if  $HK$  is a subgroup, prove the two inclusions.

To prove that  $HK$  is a subgroup if  $HK = KH$  holds, check the properties from the definition of a subgroup.

## Section 12

1. a

If  $H$  were normal then for  $a, b \in G$ , we have  $aH = Ha$  and  $bH = Hb$ . Then for two cosets  $aH, bH$ , we have

$$aHbH = abHH = (ab)H = cH$$

This shows that product of two cosets is another coset if  $H$  were normal. Now, by contraposition, if there were two cosets  $aH, bH$  such that  $aHbH$  is not any coset, then  $H$  cannot be normal.

### Result

2 of 2

Show by contraposition that if  $aHbH$  were not coset,  $H$  cannot be normal.

2. a

$H$  is a subgroup of  $G$

We check the properties from the definition of a subgroup.

Subset?

Since the determinant of each element of  $H$  is 1, which is not zero, we conclude that  $H \subseteq G$ .

Closure?

Let  $h_1, h_2 \in H$ ,

$$h_1 = \begin{bmatrix} 1 & a_1 & b_1 \\ 0 & 1 & c_1 \\ 0 & 0 & 1 \end{bmatrix}, \quad h_2 = \begin{bmatrix} 1 & a_2 & b_2 \\ 0 & 1 & c_2 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$h_1 h_2 = \begin{bmatrix} 1 & a_1 + a_2 & b_1 + b_2 + a_1 c_2 \\ 0 & 1 & c_1 + c_2 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus,  $h_1 h_2 \in H$ .

### Inverse?

Let  $h \in H$ ,

$$h = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

Now,

$$h^{-1} = \begin{bmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}$$

which can be checked by direct computation. Thus,  $h^{-1} \in H$ .

### $K$ is a normal subgroup of $H$

We check the properties from the definition of a subgroup.

#### Subset?

$K \subseteq K$  is clear.

#### Closure?

Let  $k_1, k_2 \in K$ ,

$$k_1 = \begin{bmatrix} 1 & 0 & b_1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad k_2 = \begin{bmatrix} 1 & 0 & b_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$k_1 k_2 = \begin{bmatrix} 1 & 0 & b_1 + b_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus,  $k_1 k_2 \in H$ .

#### Inverse?

Let  $k \in K$ ,

$$k = \begin{bmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now,

$$k^{-1} = \begin{bmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which can be checked by direct computation. Thus,  $k^{-1} \in K$ .

### Normal subgroup?

Let  $h \in H, k \in K$ ,

$$h = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad k = \begin{bmatrix} 1 & 0 & d \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$h^{-1} = \begin{bmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$hkh^{-1} = \begin{bmatrix} 1 & 0 & d \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in K$$

Thus,  $hkh^{-1} \in K$  for every  $h \in H, k \in K$ , so  $K$  is a normal subgroup of  $H$ .

### Quotient group $H/K$ .

Let  $h, h' \in H$  be such that

$$h = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad h' = \begin{bmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$hK = h'K$$

This is if and only if

$$h'^{-1}h \in K,$$

and since

$$h'^{-1}h = \begin{bmatrix} 1 & a-d & b-cd-e+df \\ 0 & 1 & c-f \\ 0 & 0 & 1 \end{bmatrix},$$

we conclude that  $hK = h'K$  if and only if

$$a = d, \quad c = f$$

Thus,

$$H/K = \left\{ hK \mid h = \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, a, c \in \mathbb{R} \right\}$$



### Center of $H$ .

We need to find all  $z \in H$  such that for all  $h \in H$  we have that

$$hz = zh$$

Let

$$z = \begin{bmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{bmatrix}$$

,

$$h = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

. Then

$$hz = zh$$

if and only if

$$\begin{bmatrix} 1 & a+d & b+e+af \\ 0 & 1 & c+f \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & a+d & b+cd+e \\ 0 & 1 & c+f \\ 0 & 0 & 1 \end{bmatrix}$$

which is if and only if

$$b+e+af = b+cd+e$$

This is clearly equivalent to

$$af = cd$$

The above equality must "work" with all  $a, c \in \mathbb{R}$  since  $h \in H$  was arbitrarily taken. Thus, if we set  $a = 0, c = 1$ , we get

$$d = 0$$

Similarly, setting  $a = 1, c = 0$  yields

$$f = 0$$

Thus, the only candidates for elements of a center are of the form

$$z = \begin{bmatrix} 1 & 0 & e \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad e \in \mathbb{R}$$

On the other hand, for such elements  $z$ ,

$$hz = \begin{bmatrix} 1 & a & b+e \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

$$zh = \begin{bmatrix} 1 & a & b+e \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

$$hz = zh$$

Thus, such  $z$  are in the center of  $H$ . Moreover, we can now see that  $z \in Z(H)$  (the center of  $H$ ) if and only if  $z \in K$ , so  $Z(H) = K$ .



## Result

$$H/K = \left\{ hK \mid h = \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, a, c \in \mathbb{R} \right\}$$

The center of  $H$  is  $K$ .

### 3. a

Check the properties from the definition of a subgroup first.

[Subset?](#)

It is clear that  $N \subseteq G$ .

[Closure?](#)

We need to prove that for every  $x, y \in N$ , we have that  $xy \in N$ . This is equivalent to

$$NN \subseteq N,$$

where  $NN = \{xy \mid x, y \in N\}$ .

Note that, since  $1 \in N$ , for every  $x \in N$  we have that

$$x \cdot 1 \in NN \quad \text{and} \quad x \cdot 1 = x \in N$$

So, in  $NN$  there are some elements which are in  $N$ . Now we can conclude that

$$NN \subseteq N$$

since the product set is contained entirely within one of the sets of partition, so, since there are some elements in  $N$ , all elements must be in  $N$ .

[Inverse?](#)

We need to prove that for every  $x \in N$  we have that  $x^{-1} \in N$ . Suppose that  $x^{-1} \notin N$ , and that  $x^{-1} \in M$ ,  $M \neq N$ . Then

$$xx^{-1} \in NM, \quad xx^{-1} = 1 \in N$$

So, we would have  $NM \subseteq N$  (since the product set is contained entirely within one of the sets of partition). On the other hand,

$$1x^{-1} \in NM, \quad 1x^{-1} = x^{-1} \notin N$$

So,  $NM \not\subseteq N$ , a contradiction.

Therefore, we conclude that our assumption that  $x^{-1} \notin N$  was wrong, and that we must have  $x^{-1} \in N$ .

### Normal subgroup?

We will prove that, for every  $g \in G, n \in N$ , we have that

$$gng^{-1} \in N$$

Take some  $g \in G$ . Let  $g \in A, g^{-1} \in B$  where  $A, B$  are some elements of partition  $P$ . Then we first show that

$$AN \subseteq A$$

Since  $g \in A$ ,

$$g \cdot 1 \in AN, \quad g \cdot 1 = g \in A$$

As before, since the product set  $AN$  is entirely contained in some element of partition  $P$ ,  $AN \subseteq A$ .

Now we prove that  $AB \subseteq N$ .

Since  $g \in A, g^{-1} \in B$ , we have that

$$gg^{-1} \in AB, \quad \text{and} \quad gg^{-1} = 1 \in N$$

As before,  $AB \in N$ .

Now we can prove that  $gng^{-1} \in N$  for every  $n \in N$ . To do that, first notice that  $gn \in AN \subseteq A$ , so there exists some  $a \in A$  such that  $gn = a$ . So,

$$gng^{-1} = ag^{-1} \in AB \subseteq N$$

Thus,  $gng^{-1} \in N$ . Thus,  $gng^{-1} \in N$  for every  $n \in N$ . Since we can do this for every  $g \in G$ , we conclude that  $N$  is truly a normal subgroup of  $G$ .

### Cosets?

Let  $A$  be some element of partition, and let  $a \in A$  be some element. We will prove that

$$A = aN$$

We break this equality in two inclusions.

$\subseteq$  We first prove that  $A \subseteq aN$ . Let  $b \in A$ . Let  $b^{-1} \in B$ . Then

$$b^{-1}b \in BA, \quad bb^{-1} = 1 \in N,$$

so  $BA \subseteq N$ . Specially,

$$b^{-1}a \in N,$$

so there exists some  $n \in N$  such that

$$b^{-1}a = n \iff a = bn \iff b = an^{-1}$$

Since  $N$  is a subgroup of  $G$ ,  $n^{-1} \in N$ . Thus,  $b \in aN$ , so  $A \subseteq aN$ , as required.

$\supseteq$  We now prove that  $aN \subseteq A$ . First,

$$AN \subseteq A;$$

this is because  $a \in A, 1 \in N$ , so

$$a \cdot 1 \in AN, \quad a \cdot 1 = a \in A$$

(of course, we use that the product set is contained entirely in one element of partition). Now,

$$aN = \{an \mid n \in N\} \subseteq \{an \mid a \in A, n \in N\} = AN \subseteq A$$

Thus,  $aN \subseteq A$ , as required.

Thus, we can conclude that  $A = aN$ .

## Result

4 of 4

To prove that  $N$  is a subgroup, check the properties from the definition of a subgroup. For example, for closure it can be helpful to show that  $NN \subseteq N$ .

For normality, show that  $gng^{-1} \in N$  for every  $g \in G, n \in N$ . Hint: Suppose that  $g \in A, g^{-1} \in B$ . Show that  $AN \subseteq A$  and that  $AB \subseteq N$ .

For the left part, let  $A \in P$ , and  $a \in A$ . Show that  $A = aN$ .

4. a

All cosets are of the form

$$zH = \{\pm z, \pm iz\},$$

where  $z \in \mathbb{C}^\times$ . Thus, for  $z = a + ib$ ,  $a, b \in \mathbb{R}$ , we have that

$$zH = \{a + ib, -a - ib, ia - b, b - ia\}$$

To prove that  $G/H$  is isomorphic to  $G$ , consider the function

$$\varphi : G \rightarrow G, \quad \varphi(z) = z^4$$

[ϕ homomorphism?](#)

Let  $x, y \in G$ . Then

$$\varphi(xy) = (xy)^4 = x^4 y^4 = \varphi(x)\varphi(y)$$

Thus,  $\varphi$  is a homomorphism.

[Image of ϕ?](#)

We will prove that  $\text{im}\varphi = G$ . To do this, let  $x \in G$ . Then  $x^{1/4}$  exists, it is in  $G$ , and

$$\varphi(x^{1/4}) = x$$

Thus,  $\text{im}\varphi = G$ .

[Kernel of ϕ?](#)

We need to find all  $x \in G$  such that

$$\varphi(x) = 1$$

This is equivalent to

$$x^4 = 1$$

Thus,  $x = \pm 1, \pm i$ . So,

$$\ker\varphi = H$$

[Conclusion.](#) By Theorem 2.12.10 (the First Isomorphism Theorem), we conclude that  $G/H$  is isomorphic to  $G$ .

## Result

For  $z = a + ib$ ,

$$zH = \{\pm z, \pm iz\} = \{a + ib, -a - ib, ia - b, b - ia\}$$

We can prove that  $G/H$  is isomorphic to  $G$ . Hint: consider

$$\varphi : G \rightarrow G, \quad \varphi(x) = x^4$$

and Theorem 2.12.10.

5. a

(i)

To prove that  $S$  is a subgroup of  $G$ , we check the properties from the definition of a subgroup.

Subset?  $S \subseteq G$  is trivial.

Closure?

Let  $x, y \in S$ ,

$$x = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad y = \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix}$$

Then

$$xy = \begin{bmatrix} ac & 0 \\ 0 & bd \end{bmatrix}$$

Thus,  $xy \in S$ .

Inverse?

Let  $x \in S$ ,

$$x = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

. Then we can easily check that

$$x^{-1} = \begin{bmatrix} 1/a & 0 \\ 0 & 1/b \end{bmatrix}$$

(it is well-defined since  $a \neq 0$  and  $b \neq 0$ ). So,  $x^{-1} \in S$ .

Thus,  $S$  is a subgroup of  $G$ .

### Normal subgroup?

Recall that  $S$  is a normal subgroup of  $G$  if and only if

$$gxg^{-1} \in S$$

for every  $g \in G, x \in S$ . Take

$$g = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in G \quad \text{and} \quad x = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \in S$$

Then

$$g^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

, and

$$gxg^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \notin S$$

Thus,  $S$  is **not** a normal subgroup of  $G$ .

### **(II)**

To prove that  $S$  is a subgroup of  $G$ , we check the properties from the definition of a subgroup.

Subset?  $S \subseteq G$  is trivial.

Closure?

Let  $x, y \in S$ ,

$$x = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}, \quad y = \begin{bmatrix} c & d \\ 0 & 1 \end{bmatrix}$$

Then

$$xy = \begin{bmatrix} ac & b + ad \\ 0 & 1 \end{bmatrix}$$

Thus,  $xy \in S$ .

Inverse?

Let  $x \in S$ ,

$$x = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$$

. Then we can easily check that

$$x^{-1} = \begin{bmatrix} 1/a & -b/a \\ 0 & 1 \end{bmatrix}$$

(it is well-defined since  $a \neq 0$ ). So,  $x^{-1} \in S$ .

Thus,  $S$  is a subgroup of  $G$ .

### Normal subgroup?

Recall that  $S$  is a normal subgroup of  $G$  if and only if

$$gxg^{-1} \in S$$

for every  $g \in G, x \in S$ .

Now let  $g \in G, x \in S$  be arbitrarily taken;

$$g = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \quad x = \begin{bmatrix} e & f \\ 0 & 1 \end{bmatrix}$$

Then

$$g^{-1} = \begin{bmatrix} 1/a & -b/(ad) \\ 0 & 1/d \end{bmatrix}$$

Now it is easy to see that

$$gxg^{-1} = \begin{bmatrix} e & h \\ 0 & 1 \end{bmatrix} \in S,$$

where  $h = \frac{b + af - be}{d}$ . Thus,  $S$  is a normal subgroup.

### Quotient group.

Let  $gS, hS \in G/S$  be such that  $gS = hS$ . This holds if and only if

$$h^{-1}g \in S$$

Let

$$g = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}, h = \begin{bmatrix} d & e \\ 0 & f \end{bmatrix}$$

Then

$$h^{-1} = \begin{bmatrix} 1/d & -e/(df) \\ 0 & 1/f \end{bmatrix}$$

Thus,

$$h^{-1}g = \begin{bmatrix} \frac{a}{d} & \frac{b}{d} - \frac{ce}{df} \\ 0 & \frac{c}{f} \end{bmatrix}$$

So,

$$h^{-1}g \in S$$

if and only if  $\frac{c}{f} = 1$ , which is if and only if  $c = f$ .

Thus,  $gS \neq hS$  if and only if  $c \neq f$ , which means that

$$G/S = \left\{ gS \mid g = \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}, c \in \mathbb{R} \setminus \{0\} \right\}$$

(notice that  $a, b$  do not determine whether two cosets are equal; thus, we can set  $a = 1, b = 0$ ).

(iii)

To prove that  $S$  is a subgroup of  $G$ , we check the properties from the definition of a subgroup.

Subset?  $S \subseteq G$  is trivial.

Closure?

Let  $x, y \in S$ ,

$$x = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}, \quad y = \begin{bmatrix} c & d \\ 0 & c \end{bmatrix}$$

Then

$$xy = \begin{bmatrix} ac & bc + ad \\ 0 & ac \end{bmatrix}$$

Thus,  $xy \in S$ .

Inverse?

Let  $x \in S$ ,

$$x = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$$

. Then we can easily check that

$$x^{-1} = \begin{bmatrix} 1/a & -b/a^2 \\ 0 & 1/a \end{bmatrix}$$

(it is well-defined since  $a \neq 0$ ). So,  $x^{-1} \in S$ .

Thus,  $S$  is a subgroup of  $G$ .

Normal subgroup?

Recall that  $S$  is a normal subgroup of  $G$  if and only if

$$gxg^{-1} \in S$$

for every  $g \in G, x \in S$ .

Now let  $g \in G, x \in S$  be arbitrarily taken;

$$g = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \quad x = \begin{bmatrix} e & f \\ 0 & e \end{bmatrix}$$

Then

$$g^{-1} = \begin{bmatrix} 1/a & -b/(ad) \\ 0 & 1/d \end{bmatrix}$$

Now it is easy to see that

$$gxg^{-1} = \begin{bmatrix} e & \frac{af}{d} \\ 0 & e \end{bmatrix} \in S$$

Thus,  $S$  is a normal subgroup.



### Quotient group.

Let  $gS, hS \in G/S$  be such that  $gS = hS$ . This holds if and only if

$$h^{-1}g \in S$$

Let

$$g = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}, h = \begin{bmatrix} d & e \\ 0 & f \end{bmatrix}$$

Then

$$h^{-1} = \begin{bmatrix} 1/d & -e/(df) \\ 0 & 1/f \end{bmatrix}$$

Thus,

$$h^{-1}g = \begin{bmatrix} \frac{a}{d} & \frac{b}{d} - \frac{ce}{df} \\ 0 & \frac{c}{f} \end{bmatrix}$$

So,

$$h^{-1}g \in S$$

if and only if  $\frac{c}{f} = \frac{a}{d}$ , which is if and only if  $\frac{a}{c} = \frac{d}{f}$ .

### **Result**

(I)  $S$  is a subgroup, but it is not a normal subgroup.

(II)  $S$  is a normal subgroup, and

$$G/S = \left\{ gS \mid g = \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}, c \in \mathbb{R} \setminus \{0\} \right\}$$

(III)  $S$  is a normal subgroup, and  $G/S$  is such that

$$gS = hS$$

if and only if  $\frac{a}{c} = \frac{d}{f}$ , with

$$g = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}, h = \begin{bmatrix} d & e \\ 0 & f \end{bmatrix}$$

## Miscellaneous Problem

1. a

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be an integer matrix such that its inverse  $A^{-1}$  is also an integer matrix. Since

$$\det A = ad - bc,$$

we conclude that  $\det A$  is an integer. Similarly,  $\det A^{-1}$  is an integer. Furthermore,

$$\det(I_2) = 1$$

Now,

$$\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(I_2) = 1$$

If  $\det A > 1$ , then  $\det A^{-1}$  cannot be an integer. Similarly if  $\det A < -1$ . Thus,

$$\det A = \pm 1$$

(We also get  $\det A^{-1} = \det A$ , but we do not need that here.)

If  $\det A = 1$ , then

$$ad - bc = 1 \iff ad + c(-b) = 1$$

Thus,  $a, c$  are integers such that there exist integers  $x, y$  such that

$$ax + cy = 1$$

However, this means that  $\gcd(a, c) = 1$  (since  $\gcd$  is the smallest positive integer which can be written as  $am + cn$ , for some integers  $m, n$ , and 1 is the smallest integer which exists, so we must have  $\gcd(a, c) = 1$ ).

Similarly, if  $\det A = -1$ , then

$$ad - bc = -1 \iff a(-d) + cb = 1$$

Again,  $\gcd(a, c) = 1$ .

In other words,  $a$  and  $c$  must be relatively prime.

---

## Result

2 of 2

$\gcd(a, c) = 1$ ; that is,  $a$  and  $c$  are relatively prime.

2. a

(a)

Let  $G$  be of even order. Notice that we can write

$$G = \bigcup_{g \in G} \{g, g^{-1}\}$$

However,  $1 \in G$ , and  $1^{-1} = 1$ , so  $\{1, 1^{-1}\}$  is a set of one element. If all other sets  $\{g, g^{-1}\}$  consist of two elements, then clearly  $G$  would be of odd order.

Thus, there exists some set  $\{h, h^{-1}\}$  with  $h, h^{-1} \in G$  such that it is a set of one element; meaning that

$$h = h^{-1},$$

or, after multiplying by  $h$ ,

$$h^2 = 1$$

Thus,  $h$  is an element of order 2!

(b)

Suppose that in  $G$  exists some element  $x$  of order 21. Then

$$x^{21} = 1$$

Now notice that  $y$  defined as  $y = x^7$  is of order 3.

Now suppose that in  $G$  we do not have elements of order 21. Since the order of an element divides the order of the group, the order of elements can be 1, 3, or 7. Suppose that no element is of order 3. Let  $x$  be an element of order 7. Then  $H = \langle x \rangle$  has 7 elements.

Thus, there exists some  $y \in G, y \notin H$ . It must have order 7 (because  $1 \in H$ ), so  $K = \langle y \rangle$  also has 7 elements.

Now we want to show that  $H \cap K = \{1\}$ . First of all,  $H \cap K$  is a subgroup of  $H$ . Moreover,  $y \notin H \cap K$ , so  $H \cap K$  has less elements than  $H$ .

Since the order of  $H \cap K$  divides the order of  $H$ , and  $|H| = 7$  which is prime, we conclude that  $|H \cap K| = 1$  or  $|H \cap K| = 7$ . However, it cannot be 7 because  $H \cap K$  has less elements than  $H$ . Thus,

$$|H \cap K| = 1 \implies H \cap K = \{1\}$$

Now we will show that all elements

$$x^i y^j,$$

for  $0 \leq i < 7, 0 \leq j < 7$ , are distinct. Suppose that

$$x^i y^j = x^a y^b$$

for some  $0 \leq i < 7, 0 \leq j < 7, 0 \leq a < 7, 0 \leq b < 7$ . Then

$$x^{i-a} = y^{b-j}$$

So,  $x^{i-a} \in H \cap K = \{1\}$ , so

$$x^{i-a} = 1$$

Since  $-7 < i - a < 7$ , and 7 is an order of  $x$ , we conclude that  $i - a = 0$ ; that is,  $i = a$ . Similarly, we prove that  $j = b$ .

Thus, if  $i \neq a$  or  $j \neq b$ , we have that

$$x^i y^j \neq x^a y^b$$

Thus, the set

$$\{x^i y^j \mid 0 \leq i < 7, 0 \leq j < 7\}$$

has 49 elements. But it is a subset of  $G$ , so  $G$  must have at least 49 elements! This is a contradiction since  $|G| = 21$ .

Thus,  $G$  must have at least one element of order 3.

---

## Result

3 of 3

Hint for (a):

$$G = \bigcup_{g \in G} \{g, g^{-1}\}$$

Hint for (b): Suppose that all elements are of order 7. Can you construct a subset of  $G$  which is too large?

3. a

(i)

If  $a \in G$  is of order 6, then  $\langle a \rangle$  has 6 elements. This means that  $G = \langle a \rangle$ , so  $G$  is cyclic.

Now we can prove that  $G$  is isomorphic to  $\mathbb{Z}/\mathbb{Z}6$ ; just define

$$f : G \rightarrow \mathbb{Z}/\mathbb{Z}6, \quad f(a^i) = \bar{i}$$

(ii)

Let  $b \in G$  be of order 3. Then  $N = \langle b \rangle$  is of order 3, and

$$[G : N] = \frac{|G|}{|N|} = 2$$

Thus,  $N$  is a normal subgroup of  $G$ . This means that the quotient group  $G/N$  is well-defined, and

$$|G/N| = [G : N] = 2$$

So, there exists some  $a \in G, a \notin N$ , such that

$$G/N = \{N, aN\}$$

( $a \notin N$  since if we would have  $a \in N$ , we would have  $aN = N$ ; a contradiction).

So, there are only two left cosets of  $G$ :  $N$  and  $aN$ . Since left cosets form a partition of  $G$ ,

$$G = N \cup aN$$

Now let  $g \in G$ . Then  $g \in N$  or  $g \in aN$ .

If  $g \in N$ , then  $g = b^i$ , for  $0 \leq i < 3$ .

If  $g \in aN$ , then  $g = ab^i$ , for  $0 \leq i < 3$ .

This means that

$$G = \{1, b, b^2, a, ab, ab^2\}$$

We can also see that

$$G = \langle a, b \rangle;$$

this is because  $\langle a, b \rangle \subseteq G$  clearly holds. On the other hand,

$$\{1, b, b^2, a, ab, ab^2\} \subseteq \langle a, b \rangle$$

is also clear.

One more detail which we can prove is that  $a$  is of order 2. Look at  $a^2$ . It is in  $G$ , so  $a^2 \in N$  or  $a^2 \in aN$ . If  $a^2 \in aN$ , then  $a \in N$ , which is a contradiction.

Thus,  $a^2 \in N$ . If  $a^2 = b$ , then  $a^2$  is of order 3, since  $|b| = 3$ . This means that  $a$  is of order 6; a contradiction since we assumed that  $G$  contains no element of order 6.

Similar conclusion follows if  $a^2 = b^2$ , since  $b^2$  is also of order 3.

Thus,  $a^2 = 1$ , so it is of order 2.

Now it is clear that  $G$  is isomorphic to  $S_3$ , since

$$S_3 = \{1, x, x^2, y, yx, yx^2\}$$

(the standard presentation of  $S_3$ ).

(iii)

We will first prove that  $G$  is abelian. Let  $x, y \in G$ . Then  $xy = 1$ , or  $xy$  is of order 2. If  $xy = 1$ , then  $y = x^{-1}$ , so  $xy = yx = 1$ .

If  $xy$  is of order 2, then

$$1 = (xy)^2 = xyxy$$

Multiply by  $y$  from the right and use that  $y^2 = 1$ :

$$y = xyxy^2 = xyx$$

Multiply by  $x$  from the right:

$$yx = xyx^2 = xy$$

Thus,  $G$  is abelian.

Now let  $x \in G, x \neq 1$ . Then

$$\langle x \rangle = \{1, x\}$$

Thus, there exists  $y \in G, y \notin \langle x \rangle$ . Now notice that

Now consider

$$H = \{1, x, y, xy\}$$

It is easy to see that  $H$  is a subgroup of  $G$ . It is clear that  $H \subseteq G$  and that all inverses are in  $H$  (all elements of  $H$  are of order 2; therefore, they are its own inverses). Closure is seen directly. For example,

$$y(xy) = y(yx) = y^2x = x \in H$$

Moreover,  $H$  has 4 elements; thus, there exists some  $z \in G, z \notin H$ . Consider

$$K = \{1, x, y, z, xz, xy, yz\}$$

We will prove that all elements of  $K$  are distinct. First, 1 is unique, so no other element is equal to it. Moreover,

$$x \neq y, \quad x \neq z, \quad y \neq z,$$

by definitions of  $y$  and  $z$ . Now assume that, for example,

$$x = yz$$

Then, by multiplying by  $y$  from the left,

$$z = yx = xy \in H$$

However,  $z \notin H$ ; a contradiction. Similarly, if  $xy = yz$ , then

$$yx = yz \Rightarrow x = z$$

(multiply by  $y$  from the left).

Again, a contradiction. All other combinations are checked similarly.

Thus,  $K$  has 7 elements, and  $K \subseteq G$ , since  $G$  is a group, so all products of its elements are in it. However,  $G$  has 6 elements; a contradiction!

What can we conclude from all this? That such  $G$  cannot even exist.

## Result

- (I) Such  $G$  are cyclic, and isomorphic to  $\mathbb{Z}/\mathbb{Z}6$ .
- (II) Such  $G$  are isomorphic to  $S_3$ .
- (III) Such  $G$  do not exist.

4. a

The semigroup generated by one element can be classified into two types. If  $s^i$  is distinct for all  $i \geq 1$  then the semigroup is infinite which is  $\{1, s, s^2, \dots\}$ .

If there exists integers  $m > n$  such that  $s^m = s^n$  then the semigroup is finite and consists of  $n$  elements  $\{1, s, s^2, \dots, s^{n-1}\}$ . For all  $j > n$ , there exists  $m \leq i \leq n - 1$  such that  $s^i = s^j$ .

The semigroup with  $n$  elements generated by single element may not be isomorphic to each other as  $m$  may vary from 0 to  $n - 1$ . This gives  $n$  distinct semigroup with  $n$  elements.

## Result

2 of 2

The semigroup can be finite or infinite. There exists  $n$  non isomorphic semigroups of size  $n$ .

5. a

Let  $G$  be finite semigroup generated by single element. Then  $G$  contains identity, the operation is associative with the law of composition and closure property also holds. We need only to show that if cancellation law holds then there exists inverse for every element.

Suppose  $s^m = s^n$  where  $n > m$ . Then by cancellation property, we get

$$s^m = s^m s^{n-m} \implies s^{n-m} = 1$$

This semigroup contains  $n - m$  elements and define inverse of  $s^a$  as  $s^{n-m-a}$  from which we get

$$s^{n-m-a} s^a = s^a s^{n-m-a} = s^{n-m} = 1$$

Thus inverse of every element holds hence  $G$  is a group.

## Result

2 of 2

Show that cancellation property implies inverse of every element exists.

6. a



(a)

1. Reflexive

Define function  $f(t) = a$  for all  $t \in [0, 1]$ . The function is continuous hence  $a \sim a$ .

2. Symmetric

Suppose  $a \sim b$ , then there is continuous function  $f : [0, 1] \rightarrow \mathbb{R}^k$  such that  $f(0) = a$  and  $f(1) = b$ . Define  $g(t) = f(1 - t)$  then  $g$  is also continuous function from  $[0, 1]$  to  $\mathbb{R}^k$  and  $g(0) = f(1) = b$  and  $g(1) = f(0) = a$ . Hence  $b \sim a$ .

3. Transitive

Suppose  $a \sim b$  and  $b \sim c$ . Let  $f$  and  $g$  be the paths from  $a$  to  $b$  and from  $b$  to  $c$ . Now define

$$h(t) = \begin{cases} f(2t) & \text{if } t \in [0, 1/2) \\ g(2t - 1) & \text{if } t \in [1/2, 1] \end{cases}$$

At  $t = 1/2$ ,  $h(1/2) = \lim_{t \rightarrow 1} f(t) = g(0) = b$  so it is continuous. Then this is also continuous function and since  $h(0) = f(0) = a$  and  $h(1) = g(1) = c$ , this is path from  $a$  to  $c$ . Thus  $a \sim c$ .

This shows that relation  $\sim$  defined by existence of path between two points is an equivalence relation.

(b) Relation  $\sim$  given by question defines equivalence relation on  $S$ . An equivalence relation defines a partition on subset  $S$ . Therefore  $S$  can be partitioned into subsets which are path connected i.e. if  $S_a$  is subset of  $S$  which consists of elements  $b$  such that  $b \sim a$  then every  $S_a$  is path connected since there exists path from any element or any other element.

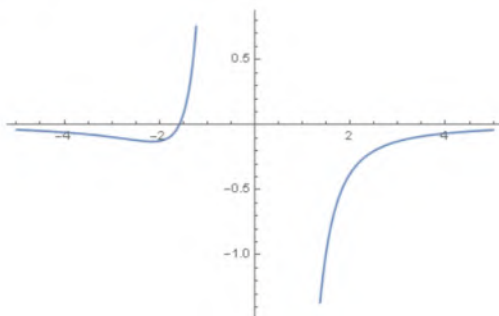
### Step 3

3 of 4

(c) The curves  $x^2 + y^2 = 1$  and  $xy = 0$  are path connected while  $xy = 1$  is not path connected. The curve  $x^2 + y^2 = 1$  is a circle so we may define  $f(t) = (\sin(ct + d), \cos(ct + d))$  where  $c$  and  $d$  are taken appropriately for any two points  $a$  and  $b$  on it.

The curve  $xy = 0$  consists of  $x = 0$  and  $y = 0$  which are pair of straight lines. They intersect at  $(0, 0)$  and since  $(0, 0)$  has path to any other point on curve. The curve is path connected.

The curve  $xy = 1$  is not path connected. Take point  $(1, 1)$  and  $(-1, -1)$ . There exists no path between these two points.



- (a) Show existence of path (continuous function) between points as required for equivalence relation. (b) Equivalence relation defines partition on set. In this case, partition is of path connected subsets. (c)  $x^2 + y^2 = 1$ ,  $xy = 0$  is path connected while  $xy = 1$  is not path connected.

7. a

- (a) Given  $G$  is subgroup of  $GL_n(\mathbb{R})$  and  $A, B, C, D \in G$ . If there exists path from  $A$  to  $B$  and from  $C$  to  $D$  then there exists path from  $AC$  to  $BD$ .

Since there exist path from  $A$  to  $C$ , let  $f$  be continuous function s.t.  $f(0) = A$  and  $f(1) = C$ . Similarly,  $g$  be continuous function such that  $g(0) = B$  and  $g(1) = D$ . Now,  $fg$  being product of continuous function is also continuous and

$$fg(0) = f(0)g(0) = AC, fg(1) = f(1)g(1) = BD$$

Thus there exists path from  $AC$  to  $BD$ .

- (b) Let  $N$  represent subset of  $G$  containing all elements which can be joined to identity matrix  $I$ . We first show that this is group.

1. Since  $I$  can be joined to itself,  $I \in N$ .
2. Suppose  $A, B \in N$  can be joined to  $I$  then by part (a), there exists path from  $AB$  to  $II = I$ . Thus closure property holds.
3. Elements of  $GL_n(\mathbb{R})$  are invertible so for any  $A$ , there exists continuous function  $\varphi_A : G \rightarrow G$  such that  $\varphi_A(X) = A^{-1}X$ . Now, suppose  $A$  can be path connected to  $I$ , then there exists continuous function  $f : [0, 1] \rightarrow GL_n(\mathbb{R})$  such that  $f(0) = I$  and  $f(1) = A$ . Now take  $g = \varphi_A \circ f$  which is continuous function such that  $g(0) = A^{-1}$  and  $g(1) = I$ . This shows  $A^{-1} \in N$ .

This shows  $N$  is subgroup of  $G$ .

To show that  $N$  is normal subgroup, let  $A \in N$ ,  $B \in G$  then we will show that  $B^{-1}AB$  is path connected to  $I$ . We know that  $A$  is path connected to  $I$ , there exists continuous function  $f$  s.t.  $f(0) = I$ . Now for any  $B \in G$ ,  $g_B(X) = B^{-1}XB$  is continuous on  $GL_n(\mathbb{R})$  so  $g = \varphi_B \circ f$  is also continuous from  $[0, 1]$  to  $GL_n(\mathbb{R})$ . Since  $g(0) = I$  and  $g(1) = B^{-1}AB$ ,  $B^{-1}AB \in N$  hence  $B^{-1}NB = N$  i.e.  $N$  is normal subgroup.

- (a) Take product function  $fg$  which forms the required path. (b) Show that  $N \subseteq G$  is subgroup first. To do that, we construct the required paths. Also make compositions with continuous functions from  $GL_n(\mathbb{R})$  to itself.

8. a

(a)

We will prove that every  $A \in SL_n(\mathbb{R})$  is connected to the identity matrix  $I_n$ .

First of all, let  $E_{ij}(\lambda)$  be some elementary matrix of the first type; it has 1 on the diagonal, and  $\lambda$  at the position  $(i, j)$ .

If  $\lambda = 0$ , then  $E_{ij}(\lambda) = I_n$ , and it is clearly connected to itself.

If  $\lambda \neq 0$ , define

$$f : [0, 1] \rightarrow SL_n(\mathbb{R}), \quad f(t) = E_{ij}(t\lambda)$$

Then  $f$  is clearly continuous (using the identification  $GL_n(\mathbb{R}) \leftrightarrow \mathbb{R}^{n \times n}$  described in the previous exercise), and  $f(0) = I_n$ ,  $f(1) = E_{ij}(\lambda)$ . Thus,  $E_{ij}(\lambda)$  is connected to  $I_n$  by a path  $f$ .

So, each elementary matrix of the first type is connected to  $I_n$ .

By the previous exercise, if  $E_1, E_2$  are two elementary matrices of the first type, then  $E_1 E_2$  is connected to  $I_n I_n = I_n$ .

By induction, now we can prove that, if  $E_1, E_2, \dots, E_k$  are elementary matrices of the first type, then  $E_1 E_2 \cdots E_k$  is connected to  $I_n$ .

Now let  $A \in SL_n(\mathbb{R})$ . Since the elementary matrices of the first type generate  $SL_n(\mathbb{R})$ , There exist elementary matrices of the first type  $F_1, \dots, F_l$  such that

$$A = F_1 \cdots F_l$$

Since  $F_1 \cdots F_l$  is connected to  $I_n$ , so is  $A$ .

Now let  $A, B \in SL_n(\mathbb{R})$  be two matrices. Since  $A$  is connected to  $I_n$ , and  $B$  is connected to  $I_n$ , and *being connected* is an equivalence relation, we conclude that  $A$  is connected to  $B$ .

Thus,  $SL_n(\mathbb{R})$  is truly path-connected.

(b)

Define

$$GL_n^+(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det A > 0\}$$

$$GL_n^-(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det A < 0\}$$

Then clearly

$$GL_n(\mathbb{R}) = GL_n^+(\mathbb{R}) \cup GL_n^-(\mathbb{R})$$

We will prove that  $GL_n^+(\mathbb{R})$  and  $GL_n^-(\mathbb{R})$  are path connected.

Let  $A \in GL_n^+(\mathbb{R})$ . Define

$$f : [0, 1] \rightarrow GL_n^+(\mathbb{R}), \quad f(t) = \frac{(1-t) + t \det A}{\det A} A$$

Since  $\det \left( \frac{(1-t) + t \det A}{\det A} A \right) > 0$  for every  $t \in [0, 1]$ ,  $f$  is well-defined. Moreover,

$$f(1) = A, \quad f(0) = A_0,$$

where  $A_1 = \frac{1}{\det A} A$ , so  $\det A_1 = 1$ , so  $A_1 \in SL_n(\mathbb{R})$ .

Now let  $A, B \in GL_n^+(\mathbb{R})$ . By the previous part,  $A$  is connected to some  $A_0 \in SL_n(\mathbb{R})$ , and  $B$  is connected to some  $B_0 \in SL_n(\mathbb{R})$ . By (a),  $A_0$  is connected to  $B_0$ . Since *being connected* is an equivalence relation, we conclude that  $A$  is connected to  $B$ . Thus,  $GL_n^+(\mathbb{R})$  is truly path-connected.

To show that  $GL_n^-(\mathbb{R})$  is path-connected, let  $A, B \in GL_n^-(\mathbb{R})$ . As with  $GL_n^+(\mathbb{R})$ . Let  $A' = AE_{1 \leftrightarrow 2}$ , where  $E_{1 \leftrightarrow 2}$  is the elementary matrix of the second kind which switches the first and the second row of  $A$ . Then  $\det A' = \det \underbrace{A \det E_{1 \leftrightarrow 2}}_{=-1} > 0$ . Define  $B'$  similarly, so  $A'$  is connected to  $B'$  by the previous part. Let

$$f : [0, 1] \rightarrow GL_n^+(\mathbb{R})$$

be the specified path. Define

$$g : [0, 1] \rightarrow GL_n^-(\mathbb{R}), \quad g(t) = f(t)E_{1 \leftrightarrow 2}$$

Firstly,

$$\det(g(t)) = \underbrace{\det(f(t))}_{>0} \underbrace{\det E_{1 \leftrightarrow 2}}_{=-1} < 0$$

Moreover,  $g$  is clearly continuous, and

$$g(0) = f(0)E_{1 \leftrightarrow 2} = A'E_{1 \leftrightarrow 2} = AE_{1 \leftrightarrow 2}E_{1 \leftrightarrow 2} = A$$

and

$$g(1) = B$$

Thus,  $GL_n^-(\mathbb{R})$  is path-connected.

To show that  $GL_n(\mathbb{R})$  is not path-connected, let  $A, B$  be such that  $\det A > 0$  and  $\det B < 0$ . If  $f$  is a path, then  $g(t) = \det f(t)$  is a continuous real function on a segment, which has positive and negative values, so for some  $t_0$  we have that  $g(t_0) = 0$ . Thus,  $f(t_0)$  is a matrix of the determinant zero, which is a contradiction, since  $f$  must have its image contained in  $GL_n(\mathbb{R})$ .

## Result

5 of 5

(a) Use the previous exercise and the hint provided with this exercise.

(b)  $GL_n(\mathbb{R}) = GL_n^+(\mathbb{R}) \cup GL_n^-(\mathbb{R})$ , where

$$GL_n^+(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det A > 0\}$$

$$GL_n^-(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det A < 0\}$$

9. a

They partition  $G$ . The proof is similar to the proof that left cosets of one subgroup partition  $G$ .

On  $G$ , define the relation  $\sim$  as

$$a \sim b \quad \text{if} \quad b = hak, \quad \text{for some } h \in H, k \in K$$

We will prove that this is an equivalence relation.

[~ reflexive?](#)

Let  $g \in G$ . Then we can write

$$g = 1g1$$

Moreover,  $1 \in H$  and  $1 \in K$ , so  $g \sim g$ .

Therefore, for every  $g \in G$  we have that  $g \sim g$ , so  $\sim$  is reflexive.

[~ symmetric?](#)

Let  $a \sim b$ . Then  $b = hak$ , for some  $h \in H, k \in K$ . Multiplying by  $h^{-1}$  from the left and  $k^{-1}$  from the right we get

$$a = h^{-1}ak^{-1}$$

Since  $H$  is a subgroup of  $G$ ,  $h^{-1} \in H$ . Similarly,  $k^{-1} \in K$ . Thus,  $b \sim a$ .

So, for every  $a, b \in G$  such that  $a \sim b$  we have that  $b \sim a$ , so  $\sim$  is symmetric.

[~ transitive?](#)

Let  $a, b, c \in G$  such that  $a \sim b, b \sim c$ . Then  $b = hak$  for some  $h \in H, k \in K$ , and  $c = h'bk'$  for some  $h' \in H, k' \in K$ . Therefore,

$$c = h'bk' = h'hakk' = (h'h)a(kk')$$

Since  $H$  is a subgroup of  $G$ ,  $h'h \in H$ . Similarly,  $kk' \in K$ . Thus,  $a \sim c$ .

So, for every  $a, b, c \in G$  such that  $a \sim b$  and  $b \sim c$  we have that  $a \sim c$ , so  $\sim$  is transitive.

Therefore, we now know that  $\sim$  is an equivalence relation. This means that its equivalence classes partition  $G$ .

Now let  $S$  be some equivalence class, and  $g \in S$ . Then

$$S = \{a \in G \mid g \sim a\} = \{a \in G \mid a = h g k, \text{ for some } h \in H, k \in K\} = \{hak \mid h \in H, k \in K\} = HaK$$

Thus, equivalence classes are equal to double cosets, so we conclude that double cosets truly partition  $G$ .

## Result

3 of 3

The proof that double cosets partition  $G$  is similar to the proof that left cosets partition  $G$ ; define a relation  $\sim$  on  $G$  with

$$a \sim b \text{ if } b = hak \text{ for some } h \in H, k \in K$$

Prove that  $\sim$  is an equivalence relations, and prove that double cosets are its equivalence classes.

10. a



$HgH = gH \implies H \text{ is normal}$

Suppose that  $HgH = gH$ , for every double coset.

Let  $g \in G, h \in H$  be arbitrarily taken. Then

$$ghg^{-1} = g(hg^{-1}h)h^{-1}$$

Since we have that

$$Hg^{-1}H = g^{-1}H,$$

and  $hg^{-1}h \in Hg^{-1}H$ , we conclude that  $hg^{-1}h \in g^{-1}H$ , so there exists some  $h' \in H$  such that  $hg^{-1}h = g^{-1}h'$ . Thus,

$$g(hg^{-1}h)h^{-1} = gg^{-1}h'h^{-1} = h'h^{-1}$$

Thus,

$$ghg^{-1} = h'h^{-1}$$

Since  $H$  is a subgroup of  $G$ , we see that  $h'h^{-1} \in H$ . Thus,  $ghg^{-1} \in H$ . Since  $g \in G$  and  $h \in H$  were taken arbitrarily, we conclude that  $H$  is a normal subgroup of  $G$  by definition.

$$\underline{H \text{ is normal} \implies HgH = gH}$$

Since  $H$  is normal, by Proposition 2.8.17. (iii),

$$gH = Hg$$

Thus,

$$HgH = gHH = g(HH)$$

(multiplication in a group is associative, so "multiplication of sets" is also associative).

Moreover, since  $H$  is a subgroup of  $G$ ,  $HH = \underbrace{\{ab \mid a, b \in H\}}_{\in H} \subseteq H$ . Thus,

$$g(HH) \subseteq gH$$

and

$$HgH \subseteq gH$$

The other inclusion is trivial; let  $x \in gH$ . Then  $x = gh$ , for some  $h \in H$ . We can also write  $x = 1gh$ , and use the fact that  $1 \in H$  to conclude that  $x \in HgH$ . Thus,  $gH \subseteq HgH$ .

Therefore, we conclude that

$$gH = HgH,$$

as required.

## Result

If  $HgH = gH$ , then let  $g \in G, h \in H$ , and show that  $ghg^{-1} \in H$ . The equality

$$ghg^{-1} = ghg^{-1}hh^{-1}$$

may prove useful.

If  $H$  is normal, then  $gH = Hg$ , so

$$HgH = g(HH)$$

Conclude the rest.

11. a

(a) Suppose we are given an invertible matrix

$$A = [a_j^i] = \begin{bmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n \end{bmatrix}$$

such that we know there is a lower triangular matrix

$$L = [l_j^i] = \begin{bmatrix} l_1^1 & 0 & \cdots & 0 \\ l_1^2 & l_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_1^n & l_2^n & \cdots & l_n^n \end{bmatrix}$$

and an upper triangular matrix with unit diagonal entries

$$U = [u_j^i] = \begin{bmatrix} 1 & u_2^1 & \cdots & u_n^1 \\ 0 & 1 & \cdots & u_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

that satisfies

$$A = LU.$$

We want to calculate the entries  $l_j^i$  and  $u_j^i$  from the entries  $a_j^i$ .

From  $A = LU$  we have by definition that for all  $1 \leq i, j \leq n$  we have

$$a_j^i = \sum_{k=1}^n l_k^i u_j^k.$$

Note that by definition  $l_j^i = 0$  if  $i < j$ ,  $u_j^i = 0$  if  $i > j$  and  $u_i^i = 1$  for all  $1 \leq i \leq n$ .

### Step 3

3 of 12

We can now proceed to calculate the columns of  $L$  and  $U$  from the first to the last.

For  $j = 1$  we have  $a_1^i = \sum_{k=1}^n l_k^i u_1^k = l_1^i$  so we have the first column of  $L$ . Note we also trivially know the first column.



Let  $j > 1$  and suppose we have already computed from the first to the  $(j - 1)$ -th columns of  $L$  and  $U$ . We want to compute  $u_j^i$  for  $i < j$  and  $l_j^i$  for  $i \geq j$ . The equations  $a_j^i = \sum_{k=1}^n l_k^i u_j^k$  then give us a system of  $n$  linear equations with  $n$  unknowns

$$\begin{cases} a_j^1 = l_1^1 u_j^1 \\ a_j^2 = l_1^2 u_j^1 + l_2^2 u_j^2 \\ \vdots \\ a_j^{j-1} = l_1^{j-1} u_j^1 + \dots + l_{j-1}^{j-1} u_j^{j-1} \\ a_j^j = l_1^j u_j^1 + \dots + l_{j-1}^j u_j^{j-1} + l_j^j u_j^j \\ a_j^{j+1} = l_1^{j+1} u_j^1 + \dots + l_{j-1}^{j+1} u_j^{j-1} + l_j^{j+1} u_j^j \\ \vdots \\ a_j^n = l_1^n u_j^1 + \dots + l_{j-1}^n u_j^{j-1} + l_j^n u_j^j \end{cases}$$

By assumption this system has a solution so any method such as Gauss elimination gives us the values for  $u_j^i$  for  $i < j$  and  $l_j^i$  for  $i \geq j$ .

### Step 5

5 of 12

After  $n$  steps of the above algorithm we have computed both  $L$  and  $U$ .

- (b) Suppose we have lower triangular matrices  $L_1$  and  $L_2$  and upper triangular matrices  $U_1$  and  $U_2$  such that  $L_1 U_1 = A = L_2 U_2$ . Note that  $\det A \neq 0$  since  $A$  is assumed to be invertible, which implies  $\det L_i \neq 0$  and  $\det U_i \neq 0$ . Therefore the  $L_i$  and  $U_i$  are all invertible.

### Step 7

7 of 12

The equation  $L_1 U_1 = L_2 U_2$  then gives us  $U_1 U_2^{-1} = L_1^{-1} L_2$ . Since the left side of the equation is upper triangular and the right side is lower triangular both sides are diagonal matrices.

### Step 8

8 of 12

Further the diagonals of  $U_1$  and  $U_2$  are composed of 1s, so the same is true for  $U_1 U_2^{-1}$  which means this is the identity matrix. We conclude that  $U_2 = U_1$ . Also  $L_1^{-1} L_2 = U_1 U_2^{-1} = Id$ , and so  $L_1 = L_2$ .

### Step 9

9 of 12

- (c) Note that for some permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  the associated permutation matrix is  $P_\sigma = [\delta_{\sigma^{-1}j}^i]$  (here  $\delta$  represents the Kronecker delta such that  $\delta_l^k = 1$  if  $k = l$  and  $\delta_l^k = 0$  if  $k \neq l$ ). Such that  $AP$  is  $A$  with its columns permuted according to the permutation  $\sigma$  and  $PA$  is  $A$  with its rows similarly permuted.

Now, for any invertible matrix  $A$  there is a lower triangular matrix  $L$  such that for each  $1 \leq j \leq n$  there is exactly one  $1 \leq i \leq n$  such that  $[A^{-1}L]_j^i = 1$  and for  $k < i$  we have  $[A^{-1}L]_j^k = 1$ . This means there is some permutation matrix  $P$  such that  $A^{-1}LP$  is upper triangular with 1s in the diagonal, which is invertible.

### Step 11

11 of 12

Defining  $U^{-1} := A^{-1}LP$  we then get  $A = LPU$ .

### Step 12

12 of 12

(d) Let  $g$  be any invertible matrix, so that any matrix of the form  $LgU$  also is invertible. By the previous item there is some permutation matrix  $P$ , lower triangular matrix  $L'$  and upper triangular with unit diagonal  $U'$  such that  $LgU = L'PU'$ , which means  $P = L'^{-1}LgUU'^{-1}$ . Therefore the  $LgU$  equals the double coset  $LPU$ .

12. a

The exercise is false when  $r, s$  must be strictly positive (you cannot write numbers of the form  $k = na$  using this method), so I assumed that  $r \geq 0, s \geq 0$ .

(a)

We will first prove that  $a$  numbers

$$0, b, 2b, (a-1)b$$

are not congruent modulo  $a$ . Suppose that

$$rb \equiv sb \text{ modulo } a,$$

with  $0 \leq r < a, 0 \leq s < b, r \neq s$ . Then  $a$  divides  $rb - sb = (r-s)b$ . However,  $a$  and  $b$  are relatively prime, so  $a$  divides  $r-s$ . Since  $-a < r-s < a$  and  $r-s \neq 0$ , this is impossible.

Thus, we have  $a$  numbers which are not even in pairs congruent modulo  $a$ . This means that, for  $k \in \mathbb{Z}$ , there exists a unique  $s \in \{0, 1, \dots, a-1\}$  such that

$$k \equiv sb \text{ modulo } a$$

Thus,  $a$  divides  $k - sb$ , so there exists some  $r \in \mathbb{Z}$  such that

$$k - sb = ra \implies k = ra + sb$$

The only problem here is that  $r \in \mathbb{Z}$ , so it does not have to be a positive integer. This means that we cannot write every integer  $k$  in a way described in the exercise. However, if  $k$  is "sufficiently large" (will be specified later), then

$$k - sb \geq 0,$$

so

$$ra \geq 0$$

Since  $a > 0$ , we must have that  $r \geq 0$ . Thus, for "sufficiently large"  $k$  the statement of the exercise holds.

(b)

Now we need to find out just how large must  $k$  be. Suppose that  $r$  is negative. Then, from

$$k = ra + sb,$$

and  $s \leq a - 1, r \leq -1$ ,

$$k \leq (a - 1)b - a = ab - a - b$$

Thus, every integer for which we have  $r < 0$  is at most  $ab - a - b$ , meaning that, if  $k > ab - a - b$ , we know that we can write it as

$$k = ra + sb$$

with  $r, s \in \mathbb{Z}, r \geq 0, s \geq 0$ .

Now we want to prove that  $ab - a - b$  cannot be written in this form, which will mean that it is also the largest positive integer which cannot be written in this form. Suppose that

$$ab - a - b = ra + sb$$

for some  $r \geq 0, s \geq 0$ . Then

$$(r + 1)a = (a - s - 1)b$$

This means that  $b$  divides  $(r + 1)a$ . On the other hand  $a$  clearly divides  $(r + 1)a$ , so  $(r + 1)a$  is a common multiple of  $a$  and  $b$ .

On the other hand, from

$$ab - a - b = ra + sb,$$

we get

$$(r + 1)a + sb = (b - 1)a$$

Since  $sb \geq 0$ , we must have that  $(r + 1)a \leq (b - 1)a$ ; thus, since  $a > 0, r + 1 \leq b - 1$ .

Putting things together now,

$$a(r + 1) \leq a(b - 1) < ab$$

Thus, there exists a common multiple of  $a$  and  $b$  which is strictly less than  $ab$ . This is however impossible since  $a$  and  $b$  are relatively prime, and their least common multiple must be  $ab$ .

Finally, we obtained a contradiction, so  $ab - a - b$  cannot be written in the way described in the exercise, which completes the proof.

## Result

4 of 4

The exercise is false when  $r, s$  must be strictly positive (you cannot write numbers of the form  $k = na$  using this method), so I assumed that  $r \geq 0, s \geq 0$ .

First prove that

$$0, b, 2b, (a - 1)b$$

are not congruent modulo  $a$ . So, for every  $k \in \mathbb{Z}$ , there exists  $s \in \{0, 1, \dots, a - 1\}$  such that  $a \equiv sb$  modulo  $a$ .

How can you now get  $r$ ? And how large must  $k$  be?

Hint: the largest positive integer which we cannot write using this method is  $k = ab - a - b$ .

Let  $S$  be the set of reachable points,  
 $T = \{(a, b) \mid a, b \in \mathbb{N}, \gcd(a, b) = 1\}$ .  
 We will prove that

$$S = T$$

We will break this equality into two inclusions.

$\subseteq$  We will first prove that  $S \subseteq T$ . To do this, let  $(x, y) \in S$ , and take the "path"

$$(1, 1) = (x_0, y_0) \rightarrow (x_1, y_1) \rightarrow \dots \rightarrow (x_n, y_n) = (x, y)$$

In each step,  $(x_n, y_n) = (x_{n-1} + y_{n-1}, y_{n-1})$  or  $(x_n, y_n) = (x_{n-1}, x_{n-1} + y_{n-1})$ . Thus, it suffices to show that if  $\gcd(a, b) = 1$ , then  $\gcd(a + b, a) = 1$  and  $\gcd(a, a + b) = 1$ .

Since  $\gcd(a, b) = 1$ , then there exist some integers  $m, n$  such that

$$ma + nb = 1$$

Now notice that

$$m(a + b) + (n - m)b = ma + mb + nb - mb = ma + nb = 1$$

Thus, there exist integers  $r, s$  such that  $r(a + b) + sb = 1$ . Since the greatest common divisor is the least positive integer which can be written in such form, and 1 is the least positive integer of them all, we conclude that  $\gcd(a + b, b) = 1$ .

The other equality,  $\gcd(a, a + b) = 1$ , is proven similarly.

Thus, if we return to our path

$$(1, 1) = (x_0, y_0) \rightarrow (x_1, y_1) \rightarrow \dots \rightarrow (x_n, y_n) = (x, y),$$

we conclude that, since  $\gcd(1, 1) = 1$ ,

$$\gcd(x_0, y_0) = 1, \gcd(x_1, y_1) = 1, \dots, \gcd(x_n, y_n) = 1, \gcd(x, y) = 1$$

Therefore,  $(x, y) \in T$ , and  $S \subseteq T$ .

$\supseteq$  Now we want to prove that  $T \subseteq S$ . That is, each point  $(a, b)$  such that  $a, b$  are positive integers and  $\gcd(a, b) = 1$  can be reached. This is done by induction.

If  $a = 1$ . Then our point is of the form  $(1, b)$ , where  $b \in \mathbb{N}$ . Now we find the path manually:

$$(1, 1) \rightarrow (1, 2) \rightarrow (1, 3) \rightarrow \dots \rightarrow (1, b)$$

Thus,  $(1, b) \in S$ .

If  $b = 1$ , then we take the path

$$(1, 1) \rightarrow (2, 1) \rightarrow (a, 1)$$

Thus,  $(a, 1) \in S$ .

Suppose that for  $(a, b) \in T$  such that  $\max\{a, b\} \leq n$  we have that  $(a, b) \in S$ , where  $n \in \mathbb{N}$ .

Now let  $(a, b) \in T$  with  $\max\{a, b\} = n + 1$ . If  $a = b$ , then  $\gcd(a, b) = \gcd(a, a) = a = n + 1 > 1$ , which is a contradiction. Thus,  $a \neq b$ . Without loss of generality, we can assume that  $a > b$ . Then  $a - b > 0$ , so  $a - b \geq 1$ . Also,  $a - b < a$ , so  $\max\{a - b, b\} < a = n + 1$ ; that is,  $\max\{a - b, b\} \leq n$ . Moreover,  $(a - b, b) \in T$  (the proof that  $\gcd(a - b, b) = 1$  is the same as in the proof of the other inclusion), so we can finally apply the Induction Hypothesis:  $(a - b, b) \in S$ . Therefore, there exists a path

$$(1, 1) = (a_0, b_0) \rightarrow (a_1, b_1) \rightarrow \dots \rightarrow (a_n, b_n) = (a - b, b)$$

Finally, we can now find a path all the way to  $(a, b)$ :

$$(1, 1) = (a_0, b_0) \rightarrow (a_1, b_1) \rightarrow \dots \rightarrow (a_n, b_n) = (a - b, b) \rightarrow (a, b)$$

Thus,  $(a, b) \in S$ , so  $T \subseteq S$ .

Conclusion. Since  $S \subseteq T$  and  $T \subseteq S$ , we now conclude that  $S = T$ .

## Result

3 of 3

Prove that

$$S = \{(a, b) \mid a, b \in \mathbb{N}, \gcd(a, b) = 1\}$$

Hint: break it up into two inclusions. One inclusion follows from  $\gcd(a, b) = \gcd(a + b, b) = \gcd(a, a + b)$ . The other inclusion can be proven by induction.

14. a



Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z})$$

. Then

$$ad - bc = 1,$$

so  $a$  and  $c$  are relatively prime. By the previous exercise, we can find a path

$$(1, 1) \rightarrow (a_1, c_1) \rightarrow \dots \rightarrow (a_n, c_n) = (a, c)$$

where  $(a_k, c_k) = (a_{k-1} + c_{k-1}, c_{k-1})$  or  $(a_k, c_k) = (a_{k-1}, a_{k-1} + c_{k-1})$ . Now notice that this means that there exist some matrices  $F_1, \dots, F_n$  where

$$F_k = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad F_k = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

such that

$$F_n \dots F_1 A = \begin{bmatrix} 1 & \tilde{b} \\ 1 & \tilde{d} \end{bmatrix} = B$$

Notice that  $\tilde{b}$  and  $\tilde{d}$  are integers.

Since clearly  $\det F_k = 1$ , we conclude that  $\det B = 1$ . Thus,

$$\tilde{d} - \tilde{b} = 1 \implies \tilde{d} = \tilde{b} + 1$$

Thus,

$$B = \begin{bmatrix} 1 & \tilde{b} \\ 1 & \tilde{b} + 1 \end{bmatrix}$$

Moreover, let

$$F = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

, then

$$FB = \begin{bmatrix} 1 & \tilde{b} \\ 0 & 1 \end{bmatrix} = C$$

Finally,

$$FF_n \dots F_1 A = FB = C$$

Notice that

$$F_k^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad F_k^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Moreover,

$$F^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

and

$$A = F_1^{-1} \dots F_n^{-1} F^{-1} C$$

All that is left is to see that

$$C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^b,$$

so  $A$  is generated by the matrices from the text of the exercise.

---

### Result

3 of 3

Use the previous exercise and the fact that the first column of matrix  $A \in SL_2(\mathbb{Z})$  consists of relatively prime integers.

15. a

!!!

16. a

The group  $\mathcal{H}$  is defined as the group generated by the letters of the alphabet subject to the relation that two words (in the group theoretical sense) that are English words with the same pronunciation are identified.

For instance  $bee = be$ , from which we can conclude by the cancellation law that  $e = 1$

---

### Step 2

2 of 6

From  $marry = merry$  we can conclude that  $a = e = 1$ .

From  $buy = by$  we can conclude that  $u = 1$ .

From  $to = too$  we get that  $o = 1$ .

From  $made = maid$  and  $a = e = 1$  we get  $md = mid$  and so that  $i = 1$ .

We conclude that all vowels are trivial in  $\mathcal{H}$ .



From *plumb* = *plum* we get  $b = 1$

From *wear* = *where* we get  $h = 1$

From *knight* = *night* we get  $k = 1$

From *mail* = *male* we get  $l = 1$

From *damn* = *dam* we get  $n = 1$

From *psalte* = *salte* we get  $p = 1$

From *fairy* = *ferry* we get  $r = 1$

From *bass* = *base* we get  $s = 1$

From *butt* = *but* we get  $t = 1$

From *chivvy* = *chivy* we get  $v = 1$

From *won* = *one* we get  $w = 1$

From *eye* = *I* we get  $y = 1$

So far the nontrivial generators of  $\mathcal{H}$  are

$$\{c, d, f, g, j, m, q, x, z\}$$

From *cell* = *sell* we get  $c = 1$

From *chased* = *chaste* we get  $d = 1$

From *right* = *write* we get  $g = 1$

From *dammed* = *damned* we get  $m = 1$

From *tax* = *tacks* we get  $x = 1$

So far the nontrivial generators of  $\mathcal{H}$  are

$$\{f, j, q, z\}$$

---

### Step 5

From *daze* = *days* we get  $z = 1$

From *phase* = *faze* we get  $f = 1$

From *genes* = *jeans* we get  $j = 1$

From *queue* = *cue* we get  $q = 1$

Therefore all generators of  $\mathcal{H}$  are trivial and so  $\mathcal{H}$  is trivial.

### Result

$\mathcal{H}$  is trivial.

# 3

## Chapter 3

### Section 1

1. a

Denote by  $S$  the specified set. We check the five properties from (3.2.1).

- Let  $x, y \in S$ . Then  $x = a + b\sqrt{2}$ ,  $y = c + d\sqrt{2}$ , for some rational numbers  $a, b, c, d$ . Now,

$$x + y = (a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$$

Since  $a + c$  and  $b + d$  are rational numbers, we conclude that  $x + y \in S$ .

- Let  $x \in S$ . Then  $x = a + b\sqrt{2}$ , for some rational numbers  $a, b$ . Moreover,  $-x = -a - b\sqrt{2}$ , since

$$(a + b\sqrt{2}) + (-a - b\sqrt{2}) = 0$$

Furthermore, we can write  $-x = (-a) + (-b)\sqrt{2}$ . Since  $-a$  and  $-b$  are rational numbers, then  $-x \in S$ .

- Let  $x, y \in S$ . Then  $x = a + b\sqrt{2}$ ,  $y = c + d\sqrt{2}$ , for some rational numbers  $a, b, c, d$ . Now,

$$xy = (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$$

Since  $(ac + 2bd)$  and  $(ad + bc)$  are rational numbers, we conclude that  $xy \in S$ .

- Let  $x \in S$ . Then  $x = a + b\sqrt{2}$ , for some rational numbers  $a, b$ . To get  $x^{-1}$ , we simplify

$$\frac{1}{x} = \frac{1}{a + b\sqrt{2}} \cdot \frac{a - b\sqrt{2}}{a - b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2} \cdot \sqrt{2}$$

Since  $\frac{a}{a^2 - 2b^2}$  and  $\frac{-b}{a^2 - 2b^2}$  are rational numbers, we conclude that  $x^{-1} = \frac{1}{x} \in S$ .

- We can write  $1 = 1 + 0\sqrt{2}$ . Since 1 and 0 are rational numbers, we get that  $1 \in S$

All properties are satisfied; hence,  $S$  is a subfield of  $\mathbb{C}$ .

#### Result

Check that properties from (3.2.1) hold.

2. a

In each case, we need to find  $\bar{x} \in \mathbb{F}_p$  such that

$$5\bar{x} = \bar{1} \iff \overline{5x} = \bar{1}$$

This is equivalent to

$$5x \equiv 1 \text{ modulo } p$$

$p = 7$

To find  $x$  such that

$$5x \equiv 1 \text{ modulo } 7,$$

we will use the Euclidean algorithm on 5 and 7:

$$\begin{aligned} 7 &= 5 \cdot 1 + 2 \implies 2 = 7 - 5 \\ 5 &= 2 \cdot 2 + 1 \implies 1 = 5 - 2 \cdot 2 \\ 2 &= 1 \cdot 2 + 0 \end{aligned}$$

Thus,  $\gcd(5, 7) = 1$ , and

$$\begin{aligned} 1 &= 5 - 2 \cdot 2 \\ &= 5 - (7 - 5) \cdot 2 \\ &= 7 \cdot (-2) + 5 \cdot 3 \end{aligned}$$

Thus,

$$5 \cdot 3 + 7 \cdot (-2) = 1$$

Since  $7 \cdot (-2) \equiv 0 \text{ modulo } 7$ , we conclude that

$$5 \cdot 3 \equiv 1 \text{ modulo } 7$$

Therefore, we take  $x = 3$ , and, considering elements in  $\mathbb{F}_7$ ,

$$\boxed{\bar{x} = \bar{3}}$$

$p = 11$

To find  $x$  such that

$$5x \equiv 1 \text{ modulo } 11,$$

we will use the Euclidean algorithm on 5 and 11:

$$\begin{aligned} 11 &= 5 \cdot 2 + 1 \implies 1 = 11 - 5 \cdot 2 \\ 5 &= 1 \cdot 5 + 0 \end{aligned}$$

Thus,  $\gcd(5, 11) = 1$ , and

$$1 = 11 - 5 \cdot 2$$

Thus,

$$5 \cdot (-2) + 11 = 1$$

Since  $11 \equiv 0 \text{ modulo } 11$ , we conclude that

$$5 \cdot (-2) \equiv 1 \text{ modulo } 11$$

Therefore, we take  $x = -2$ , and, considering elements in  $\mathbb{F}_{11}$ , and that  $\overline{-2} = \bar{9}$ ,

$$\boxed{\bar{x} = \bar{9}}$$

### $p = 13$

To find  $x$  such that

$$5x \equiv 1 \text{ modulo } 13,$$

we will use the Euclidean algorithm on 5 and 13:

$$\begin{aligned} 13 &= 5 \cdot 2 + 3 \implies 3 = 13 - 5 \cdot 2 \\ 5 &= 3 \cdot 1 + 2 \implies 2 = 5 - 3 \\ 3 &= 2 \cdot 1 + 1 \implies 1 = 3 - 2 \\ 2 &= 1 \cdot 2 + 0 \end{aligned}$$

Thus,  $\gcd(5, 13) = 1$ , and

$$\begin{aligned} 1 &= 3 - 2 \\ &= 3 - (5 - 3) \\ &= -5 + 3 \cdot 2 \\ &= -5 + (13 - 5 \cdot 2) \cdot 2 \\ &= 13 \cdot 2 + 5 \cdot (-5) \end{aligned}$$

Thus,

$$5 \cdot (-5) + 13 \cdot 2 = 1$$

Since  $13 \cdot 2 \equiv 0 \text{ modulo } 13$ , we conclude that

$$5 \cdot (-5) \equiv 1 \text{ modulo } 13$$

Therefore, we take  $x = -5$ , and, considering elements in  $\mathbb{F}_{13}$ , and that  $\overline{-5} = \overline{8}$ ,

$$\boxed{\overline{x} = \overline{8}}$$

### $p = 17$

To find  $x$  such that

$$5x \equiv 1 \text{ modulo } 17,$$

we will use the Euclidean algorithm on 5 and 17:

$$\begin{aligned} 17 &= 5 \cdot 3 + 2 \implies 2 = 17 - 5 \cdot 3 \\ 5 &= 2 \cdot 2 + 1 \implies 1 = 5 - 2 \cdot 2 \\ 2 &= 1 \cdot 2 + 0 \end{aligned}$$

Thus,  $\gcd(5, 17) = 1$ , and

$$\begin{aligned} 1 &= 5 - 2 \cdot 2 \\ &= 5 - (17 - 5 \cdot 3) \cdot 2 \\ &= 17 \cdot (-2) + 5 \cdot 7 \end{aligned}$$

Thus,

$$5 \cdot 7 + 17 \cdot (-2) = 1$$

Since  $17 \cdot (-2) \equiv 0 \text{ modulo } 17$ , we conclude that

$$5 \cdot 7 \equiv 1 \text{ modulo } 17$$

Therefore, we take  $x = 7$ , and, considering elements in  $\mathbb{F}_{17}$ ,

$$\boxed{\overline{x} = \overline{7}}$$

## Result

$$p = 7: \overline{3}$$

$$p = 11: \overline{9}$$

$$p = 13: \overline{8}$$

$$p = 17: \overline{7}$$

### 3. a

Since  $\mathbb{F}_7$  is commutative,

$$(ax^\alpha)(bx^\beta) = abx^{\alpha+\beta}$$

Moreover, since we can apply the Distributive Property, the starting product is the same of the following polynomials:

$$\begin{aligned} & (x^3 + 3x^2 + 3x + 1)x^4 \\ & (x^3 + 3x^2 + 3x + 1)(4x^3) \\ & (x^3 + 3x^2 + 3x + 1)(6x^2) \\ & (x^3 + 3x^2 + 3x + 1)(4x) \\ & (x^3 + 3x^2 + 3x + 1)1 \end{aligned}$$

Applying the Distributive Property again:

$$\begin{aligned} (x^3 + 3x^2 + 3x + 1)x^4 &= x^3 \cdot x^4 + (3x^2)x^4 + (3x)x^4 + 1x^4 \\ &= x^7 + 3x^6 + 3x^5 + x^4 \\ (x^3 + 3x^2 + 3x + 1)(4x^3) &= x^3(4x^3) + (3x^2)(4x^3) + (3x)(4x^3) + 1(4x^3) \\ &= 4x^6 + 12x^5 + 12x^4 + 4x^3 \\ &= 4x^6 + 5x^5 + 5x^4 + 4x^3 \end{aligned}$$

(The last equality follows because  $12 = 5$  in  $\mathbb{F}_7$ .)

$$\begin{aligned} (x^3 + 3x^2 + 3x + 1)(6x^2) &= x^3(6x^2) + (3x^2)(6x^2) + (3x)(6x^2) + 1(6x^2) \\ &= 6x^5 + 18x^4 + 18x^3 + 6x^2 \\ &= 6x^5 + 4x^4 + 4x^3 + 6x^2 \end{aligned}$$

(The last equality follows because  $18 = 4$  in  $\mathbb{F}_7$ .)

$$\begin{aligned} (x^3 + 3x^2 + 3x + 1)(4x) &= x^3(4x) + (3x^2)(4x) + (3x)(4x) + 1(4x) \\ &= 4x^4 + 12x^3 + 12x^2 + 4x \\ &= 4x^4 + 5x^3 + 5x^2 + 4x \end{aligned}$$

(The last equality follows because  $12 = 5$  in  $\mathbb{F}_7$ .)

The equality

$$(x^3 + 3x^2 + 3x + 1)1 = x^3 + 3x^2 + 3x + 1$$

is trivial.

Now the original product is equal to the sum of all these polynomials; so, it is equal to

$$(x^7 + 3x^6 + 3x^5 + x^4) + (4x^6 + 5x^5 + 5x^4 + 4x^3) + (6x^5 + 4x^4 + 4x^3 + 6x^2) + (4x^4 + 5x^3 + 5x^2 + 4x) + (x^3 + 3x^2 + 3x + 1)$$

Using the Distributive Property on  $x^c$ , this is equal to

$$\begin{aligned} & x^7 + (3+4)x^6 + (3+5+6)x^5 + (1+5+4+4)x^4 + (4+4+5+1)x^3 + (6+5+3)x^2 + (4+3)x + 1 \\ &= x^7 + 7x^6 + 14x^5 + 14x^4 + 14x^3 + 14x^2 + 7x + 1 \\ &= \boxed{x^7 + 1} \end{aligned}$$

since  $7 = 14 = 0$  in  $\mathbb{F}_7$ .

## Result

$$x^7 + 1$$

4. a

Let

$$A = \begin{bmatrix} 6 & -3 \\ 2 & 6 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\underline{p = 5}$$

Firstly,

$$\det A = 36 + 3 \cdot 2 = 42 = 2$$

(the last equality holds in  $\mathbb{F}_5$ ).

Thus,  $\det A \neq 0$ . This means that  $A^{-1}$  exists; we use the following formula:

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C^{-1} = (\det C)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Furthermore,  $2^{-1} = 3$  in  $\mathbb{F}_5$  (since  $2 \cdot 3 = 3 \cdot 2 = 6 = 1$ ).

Thus,

$$A^{-1} = 3 \cdot \begin{bmatrix} 6 & 3 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 18 & 9 \\ -6 & 18 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix}$$

Thus,

$$X = A^{-1}B = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 13 \\ 15 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

To write it clearly,

$$\boxed{X = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \text{ modulo } 5}$$

### $p = 11$

Firstly,

$$\det A = 36 + 3 \cdot 2 = 42 = 9$$

(the last equality holds in  $\mathbb{F}_{11}$ ).

Thus,  $\det A \neq 0$ . This means that  $A^{-1}$  exists; we use the following formula:

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C^{-1} = (\det C)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Furthermore,  $9^{-1} = 5$  in  $\mathbb{F}_{11}$  (since  $9 \cdot 5 = 5 \cdot 9 = 45 = 1$ ).

Thus,

$$A^{-1} = 5 \cdot \begin{bmatrix} 6 & 3 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 30 & 15 \\ -10 & 30 \end{bmatrix} = \begin{bmatrix} 8 & 4 \\ 1 & 8 \end{bmatrix}$$

Thus,

$$X = A^{-1}B = \begin{bmatrix} 8 & 4 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 28 \\ 11 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

To write it clearly,

$$X = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \text{ modulo } 11$$

### $p = 13$

Firstly,

$$\det A = 36 + 3 \cdot 2 = 42 = 3$$

(the last equality holds in  $\mathbb{F}_{13}$ ).

Thus,  $\det A \neq 0$ . This means that  $A^{-1}$  exists; we use the following formula:

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C^{-1} = (\det C)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Furthermore,  $3^{-1} = 9$  in  $\mathbb{F}_{13}$  (since  $3 \cdot 9 = 9 \cdot 3 = 27 = 1$ ).

Thus,

$$A^{-1} = 9 \cdot \begin{bmatrix} 6 & 3 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 54 & 27 \\ -18 & 54 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 8 & 2 \end{bmatrix}$$

Thus,

$$X = A^{-1}B = \begin{bmatrix} 2 & 1 \\ 8 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 26 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \end{bmatrix}$$

To write it clearly,

$$X = \begin{bmatrix} 7 \\ 0 \end{bmatrix} \text{ modulo } 13$$



### $p = 17$

Firstly,

$$\det A = 36 + 3 \cdot 2 = 42 = 8$$

(the last equality holds in  $\mathbb{F}_{17}$ ).

Thus,  $\det A \neq 0$ . This means that  $A^{-1}$  exists; we use the following formula:

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C^{-1} = (\det C)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Furthermore,  $8^{-1} = 15$  in  $\mathbb{F}_{17}$  (since  $8 \cdot 15 = 15 \cdot 8 = 120 = 1$ ).

Thus,

$$A^{-1} = 15 \cdot \begin{bmatrix} 6 & 3 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 90 & 45 \\ -30 & 90 \end{bmatrix} = \begin{bmatrix} 5 & 11 \\ 4 & 5 \end{bmatrix}$$

Thus,

$$X = A^{-1}B = \begin{bmatrix} 5 & 11 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 26 \\ 17 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

To write it clearly,

$$X = \begin{bmatrix} 9 \\ 0 \end{bmatrix} \text{ modulo } 17$$

### $p = 7$

Firstly,

$$\det A = 36 + 3 \cdot 2 = 42 = 0$$

(the last equality holds in  $\mathbb{F}_7$ ).

Thus,  $\det A = 0$ . We therefore cannot find the inverse  $A^{-1}$ . We solve this system a bit differently.

First,

$$A = \begin{bmatrix} -1 & -3 \\ 2 & -1 \end{bmatrix}$$

The system

$$AX = B$$

can be written as two equalities:

$$\begin{aligned} -x_1 - 3x_2 &= 3 \\ 2x_1 - x_2 &= 1 \end{aligned}$$

Multiply the first equation by 2 and add it to the second equation:

$$\begin{aligned} -x_1 - 3x_2 &= 3 \\ -7x_2 &= 1 \end{aligned}$$

However,  $-7x_2 = 0$  in  $\mathbb{F}_7$ , so the second equation becomes  $0 = 1$ , which clearly does not hold in  $\mathbb{F}_7$ . Therefore, this system has **no solutions**.

## Result

$$p = 5 : X = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \text{ modulo } 5$$

$$p = 11 : X = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \text{ modulo } 11$$

$$p = 13 : X = \begin{bmatrix} 7 \\ 0 \end{bmatrix} \text{ modulo } 13$$

$$p = 17 : X = \begin{bmatrix} 9 \\ 0 \end{bmatrix} \text{ modulo } 17$$

$$p = 7 : \text{no solutions}$$

5. a

This matrix is invertible if and only if  $\det A \neq 0$ . We will first find the matrix in  $\mathbb{R}$  using the following row operation: multiply the  $j$ th row with the constant  $c$  and add it to the  $i$ th row:  $\xrightarrow{cR_j+R_i}$ .

$$\left| \begin{array}{ccc} 1 & 2 & 0 \\ 0 & 3 & -1 \\ -2 & 0 & 2 \end{array} \right| \xrightarrow{2R_1+R_2} \left| \begin{array}{ccc} 1 & 2 & 0 \\ 0 & 3 & -1 \\ 0 & 4 & 2 \end{array} \right|$$

Now apply the LaPlace Expansion on the first column:

$$\left| \begin{array}{ccc} 1 & 2 & 0 \\ 0 & 3 & -1 \\ 0 & 4 & 2 \end{array} \right| = \left| \begin{array}{cc} 3 & -1 \\ 4 & 2 \end{array} \right| = 10$$

Thus,

$$\det A = 10$$

Now  $\det A = 0$  in  $\mathbb{F}_p$  if and only if  $p$  divides 10. Since  $10 = 2 \cdot 5$ , we get that  $\det A = 0$  in  $\mathbb{F}_p$  if and only if  $p = 2$  or  $p = 5$ .

Thus,  $A$  has an inverse (equivalently,  $\det A \neq 0$ ) if and only if  $p \neq 2$  and  $p \neq 5$ .

## Result

$$p \neq 2 \text{ and } p \neq 5.$$

6. a

We always use

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Solving in  $\mathbb{Q}$ .

System  $AX = 0$ .

Here  $AX = 0$  is equivalent to the system of equations

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 &+ x_3 = 0 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 0$$

Thus,  $x_1 = 0$ . Moreover, we now easily obtain  $x_2 = 0$  and  $x_3 = 0$ . Thus,

$$X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

System  $AX = B$ .

Now onto the system  $AX = B$ . This becomes

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 &+ x_3 = -1 \\ x_1 - x_2 - x_3 &= 1 \end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 1$$

Therefore,  $x_1 = \frac{1}{3}$ . Now, from the first and the second equation, we can easily obtain  $x_2 = \frac{2}{3}$  and  $x_3 = -\frac{4}{3}$ .

Thus,

$$X = \begin{bmatrix} 1/3 \\ 2/3 \\ -4/3 \end{bmatrix}$$

Solving in  $\mathbb{F}_2$ .

System  $AX = 0$ .

Here  $AX = 0$  is equivalent to the system of equations

$$\begin{aligned}x_1 + x_2 &= 0 \\x_1 &+ x_3 = 0 \\x_1 - x_2 - x_3 &= 0\end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 0$$

Moreover,  $3 = 1$  in  $\mathbb{F}_2$ , so  $3x_1 = x_1$  and  $x_1 = 0$ . Moreover, we now easily obtain  $x_2 = 0$  and  $x_3 = 0$ . Thus,

$$X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (\text{modulo } 2)$$

System  $AX = B$ .

Now onto the system  $AX = B$ . This becomes

$$\begin{aligned}x_1 + x_2 &= 1 \\x_1 &+ x_3 = -1 \\x_1 - x_2 - x_3 &= 1\end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 1$$

Therefore, since  $3x_1 = x_1$ ,  $x_1 = 1$ . Now, from the first and the second equation, we can easily obtain  $x_2 = 0$  and  $x_3 = -2 = 0$  (because  $-2 = 0$  in  $\mathbb{F}_2$ ). Thus,

$$X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (\text{modulo } 2)$$

### Solving in $\mathbb{F}_3$ .

#### System $AX = 0$ .

Here  $AX = 0$  is equivalent to the system of equations

$$\begin{aligned}x_1 + x_2 &= 0 \\x_1 &+ x_3 = 0 \\x_1 - x_2 - x_3 &= 0\end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 0$$

Moreover,  $3 = 0$  in  $\mathbb{F}_3$ , so  $3x_1 = 0$ . Thus, the third equation is  $0 = 0$ . The whole system is

$$\begin{aligned}x_1 + x_2 &= 0 \\x_1 &+ x_3 = 0 \\0 &= 0\end{aligned}$$

Thus,

$$\begin{aligned}x_2 &= -x_1 = 2x_1 \\x_3 &= -x_1 = 2x_1\end{aligned}$$

(since  $-1 = 2$  in  $\mathbb{F}_3$ , it follows that  $-x_1 = 2x_1$ ). So, when we set  $t = x_1$  as a parameter,

$$X = \begin{bmatrix} t \\ 2t \\ 2t \end{bmatrix}, t \in \mathbb{F}_3 \quad (\text{modulo } 3)$$

#### System $AX = B$ .

Now onto the system  $AX = B$ . This becomes

$$\begin{aligned}x_1 + x_2 &= 1 \\x_1 &+ x_3 = -1 \\x_1 - x_2 - x_3 &= 1\end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 1$$

Therefore, since  $3x_1 = 0$ , the third equation is  $0 = 1$ . This is however absurd since  $0 \neq 1$  in  $\mathbb{F}_3$ . Thus, this system has **no solution**.

Solving in  $\mathbb{F}_7$ .

System  $AX = 0$ .

Here  $AX = 0$  is equivalent to the system of equations

$$\begin{aligned}x_1 + x_2 &= 0 \\x_1 &+ x_3 = 0 \\x_1 - x_2 - x_3 &= 0\end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 0$$

Since  $\mathbb{F}_7$  is a field, the inverse of 3,  $3^{-1}$ , exists. Multiplying the above equation by  $3^{-1}$ ,

$$3^{-1} \cdot 3x_1 = 3^{-1} \cdot 0 \implies \boxed{x_1 = 0}$$

Now from the first and the second equation, we now easily obtain

$$\boxed{x_2 = 0}, \quad \boxed{x_3 = 0}$$

Thus,

$$X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (\text{modulo } 7)$$

System  $AX = B$ .

Now onto the system  $AX = B$ . This becomes

$$\begin{aligned}x_1 + x_2 &= 1 \\x_1 &+ x_3 = -1 \\x_1 - x_2 - x_3 &= 1\end{aligned}$$

Adding the first and the second equation to the third, the third equation becomes

$$3x_1 = 1$$

Now we want to find  $3^{-1}$  explicitly. Notice that  $3 \cdot 5 = 5 \cdot 3 = 15 = 1$  (in  $\mathbb{F}_7$ ), so  $3^{-1} = 5$ . So, multiplying the equation  $3x_1 = 1$  by 5,

$$15x_1 = 5 \implies \boxed{x_1 = 5}$$

The first equation is now

$$5 + x_2 = 1 \quad / \quad -5 \implies x_2 = -4 = 3 \implies \boxed{x_2 = 3}$$

Similarly, the second is

$$5 + x_3 = -1 \quad / \quad -5 \implies x_3 = -6 = 1 \implies \boxed{x_3 = 1}$$

Therefore,

$$X = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix} \quad (\text{modulo } 7)$$

## Result

$\mathbb{Q}$ :  $AX = 0$  has a solution

$$X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

,  $AX = B$  has a solution

$$X = \begin{bmatrix} 1/3 \\ 2/3 \\ -4/3 \end{bmatrix}$$

$\mathbb{F}_2$ :  $AX = 0$  has a solution

$$X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

,  $AX = B$  has a solution

$$X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$\mathbb{F}_3$ :  $AX = 0$  has a solution

$$X = \begin{bmatrix} t \\ 2t \\ 2t \end{bmatrix}$$

,  $t \in \mathbb{F}_3$ ,  $AX = B$  has no solution.

$\mathbb{F}_7$ :  $AX = 0$  has a solution

$$X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

,  $AX = B$  has a solution

$$X = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$$

7. a



### $p = 2$

In this case,

$$F_2 = \{0, 1\}$$

Therefore,

$$F_2^\times = \{1\},$$

which is clearly cyclic with the generator 1.

### $p = 3$

In this case,

$$F_3 = \{0, 1, 2\}$$

Therefore,

$$F_3^\times = \{1, 2\},$$

Since  $2^2 = 4 = 1$ , this group is cyclic with the generator 2.

### $p = 5$

In this case,

$$F_5 = \{0, 1, 2, 3, 4\}$$

Therefore,

$$F_5^\times = \{1, 2, 3, 4\},$$

Calculate potentions of 2 up to the third power ( $p - 2$ ):

$$2^2 = 4$$

$$2^3 = 8 = 3$$

Therefore,

$$\{2^0, 2^1, 2^2, 2^3\} = \{1, 2, 4, 3\} = F_5^\times,$$

and this group is cyclic with the generator 2.

### $p = 7$

In this case,

$$F_7 = \{0, 1, 2, 3, 4, 5, 6\}$$

Therefore,

$$F_7^\times = \{1, 2, 3, 4, 5, 6\},$$

Calculate potentions of 3 up to the 5th power ( $p - 2$ ):

$$3^2 = 9 = 2$$

$$3^3 = 3^2 \cdot 3 = 2 \cdot 3 = 6$$

$$3^4 = 3^3 \cdot 3 = 6 \cdot 3 = 18 = 4$$

$$3^5 = 3^4 \cdot 3 = 4 \cdot 3 = 12 = 5$$

Therefore,

$$\{3^0, 3^1, 3^2, 3^3, 3^4, 3^5\} = \{1, 3, 2, 6, 4, 5\} = F_7^\times,$$

and this group is cyclic with the generator 3.

### $p = 11$

In this case,

$$F_{11} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Therefore,

$$F_{11}^\times = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\},$$

Calculate potentions of 2 up to the 9th power ( $p - 2$ ):

$$2^2 = 4$$

$$2^3 = 8$$

$$2^4 = 16 = 5$$

$$2^5 = 2^4 \cdot 2 = 5 \cdot 2 = 10$$

$$2^6 = 2^5 \cdot 2 = 10 \cdot 2 = 20 = 9$$

$$2^7 = 2^6 \cdot 2 = 9 \cdot 2 = 18 = 7$$

$$2^8 = 2^7 \cdot 2 = 7 \cdot 2 = 14 = 3$$

$$2^9 = 2^8 \cdot 2 = 3 \cdot 2 = 6$$

Therefore,

$$\{2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9\} = \{1, 2, 4, 8, 5, 10, 9, 7, 3, 6\} = F_{11}^\times,$$

and this group is cyclic with the generator 2.

### $p = 13$

In this case,

$$F_{13} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Therefore,

$$F_{13}^{\times} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\},$$

Calculate potentions of 2 up to the 11th power ( $p - 2$ ):

$$2^2 = 4$$

$$2^3 = 8$$

$$2^4 = 16 = 3$$

$$2^5 = 2^4 \cdot 2 = 3 \cdot 2 = 6$$

$$2^6 = 2^5 \cdot 2 = 6 \cdot 2 = 12$$

$$2^7 = 2^6 \cdot 2 = 12 \cdot 2 = 24 = 11$$

$$2^8 = 2^7 \cdot 2 = 11 \cdot 2 = 22 = 9$$

$$2^9 = 2^8 \cdot 2 = 9 \cdot 2 = 18 = 5$$

$$2^{10} = 2^9 \cdot 2 = 5 \cdot 2 = 10$$

$$2^{11} = 2^{10} \cdot 2 = 10 \cdot 2 = 20 = 7$$

Therefore,

$$\{2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}, 2^{11}\} = \{1, 2, 4, 8, 3, 6, 12, 11, 9, 5, 10, 7\} = F_{13}^{\times},$$

and this group is cyclic with the generator 2.

### $p = 17$

In this case,

$$F_{17} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16\}$$

Therefore,

$$F_{17}^{\times} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16\},$$

Calculate potentions of 3 up to the 15th power ( $p - 2$ ):

$$3^2 = 9$$

$$3^3 = 27 = 10$$

$$3^4 = 3^3 \cdot 3 = 10 \cdot 3 = 30 = 13$$

$$3^5 = 3^4 \cdot 3 = 13 \cdot 3 = 39 = 5$$

$$3^6 = 3^5 \cdot 3 = 5 \cdot 3 = 15$$

$$3^7 = 3^6 \cdot 3 = 15 \cdot 3 = 45 = 11$$

$$3^8 = 3^7 \cdot 3 = 11 \cdot 3 = 33 = 16$$

$$3^9 = 3^8 \cdot 3 = 16 \cdot 3 = 48 = 14$$

$$3^{10} = 3^9 \cdot 3 = 14 \cdot 3 = 42 = 8$$

$$3^{11} = 3^{10} \cdot 3 = 8 \cdot 3 = 24 = 7$$

$$3^{12} = 3^{11} \cdot 3 = 7 \cdot 3 = 21 = 4$$

$$3^{13} = 3^{12} \cdot 3 = 4 \cdot 3 = 12$$

$$3^{14} = 3^{13} \cdot 3 = 12 \cdot 3 = 36 = 2$$

$$3^{15} = 3^{14} \cdot 3 = 2 \cdot 3 = 6$$

Therefore,

$$\begin{aligned} & \{3^0, 3^1, 3^2, 3^3, 3^4, 3^5, 3^6, 3^7, 3^8, 3^9, 3^{10}, 3^{11}, 3^{12}, 3^{13}, 3^{14}, 3^{15}\} \\ &= \{1, 3, 9, 10, 13, 5, 15, 11, 16, 14, 8, 7, 4, 12, 2, 6\} \\ &= F_{17}^\times, \end{aligned}$$

and this group is cyclic with the generator 3.

**\underline{p = 19}**

In this case,

$$F_{19} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19\}$$

Therefore,

$$F_{19}^\times = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19\},$$

Calculate potensions of 2 up to the 17th power ( $p - 2$ ):

$$2^2 = 4$$

$$2^3 = 8$$

$$2^4 = 16$$

$$2^5 = 32 = 13$$

$$2^6 = 2^5 \cdot 2 = 13 \cdot 2 = 26 = 7$$

$$2^7 = 2^6 \cdot 2 = 7 \cdot 2 = 14$$

$$2^8 = 2^7 \cdot 2 = 14 \cdot 2 = 28 = 9$$

$$2^9 = 2^8 \cdot 2 = 9 \cdot 2 = 18$$

$$2^{10} = 2^9 \cdot 2 = 18 \cdot 2 = 36 = 17$$

$$2^{11} = 2^{10} \cdot 2 = 17 \cdot 2 = 34 = 15$$

$$2^{12} = 2^{11} \cdot 2 = 15 \cdot 2 = 30 = 11$$

$$2^{13} = 2^{12} \cdot 2 = 11 \cdot 2 = 22 = 3$$

$$2^{14} = 2^{13} \cdot 2 = 3 \cdot 2 = 6$$

$$2^{15} = 2^{14} \cdot 2 = 6 \cdot 2 = 12$$

$$2^{16} = 2^{15} \cdot 2 = 12 \cdot 2 = 24 = 5$$

$$2^{17} = 2^{16} \cdot 2 = 5 \cdot 2 = 10$$

Therefore,

$$\begin{aligned} & \{2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}, 2^{11}, 2^{12}, 2^{13}, 2^{14}, 2^{15}, 2^{16}, 2^{17}\} \\ &= \{1, 2, 4, 8, 16, 13, 7, 14, 9, 18, 17, 15, 11, 3, 6, 12, 5, 10\} \\ &= F_{19}^\times, \end{aligned}$$

and this group is cyclic with the generator 2.

## Result

Some primitive roots are as follows:

$$p = 2: 1$$

$$p = 3: 2$$

$$p = 5: 2$$

$$p = 7: 3$$

$$p = 11: 2$$

$$p = 13: 2$$

$$p = 17: 3$$

$$p = 19: 2$$

8. a

(a)

If  $p$  divides  $a$ , then the result is trivial, since

$$a \equiv 0 \quad \text{and} \quad a^p \equiv 0 \quad \text{modulo } p$$

If  $p$  does not divide  $a$ , then  $\bar{a} \in \mathbb{F}_p^\times$ . Since this is a group of order  $p - 1$ , we conclude that

$$\bar{a}^{p-1} = \bar{1}$$

This is because the order of  $\bar{a}$  divides the order of the group; so,  $p - 1 = k|\bar{a}|$ , for some positive integer  $k$ , where  $|\bar{a}|$  is the order of  $\bar{a}$ . Now,

$$\bar{a}^{p-1} = (\bar{a}^{|\bar{a}|})^k = \bar{1}^k = \bar{1}$$

From  $\bar{a}^{p-1} = \bar{1}$ , by multiplying by  $\bar{a}$ , we get

$$\bar{a}^p = \bar{a}$$

This means that

$$a^p \equiv a \quad \text{modulo } p,$$

as required.

(b)

If  $p = 2$ , then  $(p - 1)! = 1 \equiv -1 \quad \text{modulo } 2$ .

If  $p = 3$ , then  $(p - 1)! = 2 \equiv -1 \quad \text{modulo } 3$ .

Now let  $p > 3$ . Since  $\mathbb{F}_p^\times$  is a field, each  $\bar{a} \in \mathbb{F}_p^\times = \{\bar{1}, \dots, \overline{p-1}\}$  has a unique inverse  $\bar{b}$  (we also assume that  $0 < a < p$ ). Suppose that  $\bar{b} = \bar{a}$ . Then

$$\bar{a}\bar{b} = \bar{1} \iff \overline{a^2 - 1} = \bar{0}$$

This means that  $p$  divides  $a^2 - 1 = (a - 1)(a + 1)$ . Since  $p$  is prime,  $p$  divides one of the factors.

If  $p$  divides  $a - 1$ , then, because  $0 \leq a - 1 < p$ , we must have that  $a - 1 = 0$ ; that is,  $a = 1$ , and  $\bar{a} = \bar{1}$ .

If  $p$  divides  $a + 1$ , then, because  $0 < a + 1 \leq p$ , we must have that  $a + 1 = p$ ; that is,  $a = p - 1$ , and  $\bar{a} = \overline{p-1}$ .

All other elements have an inverse which is different from themselves; thus, we can pair them up (because  $\bar{a}$  has a unique inverse  $\bar{b} \neq \bar{a}$ , and  $\bar{a}$  is also a unique inverse of  $\bar{b}$ ), from which we get

$$\bar{2} \cdot \bar{3} \cdots \overline{p-2} = \bar{1}$$

Thus,

$$\overline{(p-1)!} = \overline{p-1} \cdot \bar{1} = \overline{p-1} = -\bar{1}$$

Therefore,

$$\overline{(p-1)!} = -\bar{1},$$

which by definition means that

$$(p-1)! \equiv -1 \quad \text{modulo } p,$$

as required.

(a) If  $p$  divides  $a$ , the statement is trivial. If it does not, then  $\bar{a} \in \mathbb{F}_p^\times$ . What is the order of this group? How can it help you to prove the rest?

(b) Each element  $\bar{a} \in \mathbb{Z}_p^\times$  has a unique inverse. Show that

$$2 \cdot 3 \cdots (p-2) \equiv 1 \text{ modulo } p$$

Prove the rest.

9. a

Notice that the identity element is the identity matrix

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

. Then

$$A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$$

$$A^5 = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$$

$$A^6 = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}$$

$$A^7 = \begin{bmatrix} 1 & 7 \\ 0 & 1 \end{bmatrix}$$



Now we use that  $7 = 0$  in  $\mathbb{F}_7$  to conclude that

$$A^7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, the order of  $A$  is 7.

Now let

$$B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

. Then

$$B^2 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

$$B^3 = \begin{bmatrix} 8 & 0 \\ 0 & 1 \end{bmatrix}$$

Now we use that  $8 = 1$  in  $\mathbb{F}_7$  to conclude that

$$B^3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, the order of  $B$  is 3.

## Result

The order of the first matrix is 7, the order of the second matrix is 3.

## 10. a

Let  $S$  be the described set, and

$$A_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Addition makes  $S$  into an abelian group.

Associativity. This follows from the fact that matrix addition is associative.

Identity. It is enough to notice that

$$A_0 + A_i = A_i + A_0 = A_i$$

for all  $i = 0, 1, 2, 3$ . Thus,  $A_0$  is the identity.

Inverse. Notice that

$$A_i + A_i = A_0,$$

since all elements of  $A_i + A_i$  will be either 0 or 2. If they are 2, we have that  $2 = 0$  in  $\mathbb{F}_2$ . Thus, the equality holds.

Commutativity. This follows from the fact that matrix addition is commutative.

Now first note that  $S^\times = \{A_1, A_2, A_3\}$ .

Multiplication makes  $S^*$  into a group.

Associativity. This follows from the fact that matrix multiplication is associative.

Identity. It is enough to notice that

$$A_1 A_i = A_i A_1$$

for all  $i = 1, 2, 3$ . Thus,  $A_1$  is the identity.

Inverse. We will find the inverse of each element. First of all,

$$A_1^2 = A_1$$

is trivial. Moreover,

$$A_2 A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A_1$$

$$A_3 A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A_1$$

Therefore,  $A_2^{-1} = A_3$  and  $A_3^{-1} = A_2$ , which proves that each element has an inverse.

Multiplication on  $S$  is commutative.

First,

$$A_0 A_i = A_0 = A_i A_0$$

for all  $i = 0, 1, 2, 3$ , so  $A_0$  commutes with each element.

Moreover,  $A_1$  is the identity in  $S^\times$ , so

$$A_1 A_i = A_i = A_i A_1$$

for  $i = 1, 2, 3$ , so it also commutes with each element (that it commutes with  $A_0$  was checked before).

Finally,

$$A_2 A_3 = A_1 = A_3 A_2$$

Now we conclude that

$$A_i A_j = A_j A_i$$

for all  $i, j \in \{0, 1, 2, 3\}$ , so multiplication is commutative.

Distributive Property.

This holds since it holds for the standard matrix operations.

## Result

Denote the set of this matrices by  $S$ , and the set of its nonzero elements by  $S^\times$ . Check that:

$(S, +)$  is an abelian group.

$(S^\times, \cdot)$  is a group.

Multiplication is commutative.

The distributive property holds.

## 11. a

Let  $S$  be the described set.

Addition makes  $S$  into an abelian group.

Associativity. This follows from the fact that addition of complex numbers is associative.

Identity. It is enough to notice that  $0 \in S$ . Since  $0$  is the identity in  $\mathbb{C}$ , it will also be the identity here.

Inverse. Let  $a + bi \in S$ . Then  $a \in \mathbb{F}_3$ , so there exists its additive inverse  $-a \in \mathbb{F}_3$ . Similarly,  $-b \in \mathbb{F}_3$  exists. Therefore,  $-a + (-b)i \in S$ . Now,

$$(a + bi) + (-a + (-b)i) = (a + (-a)) + (b + (-b))i = 0$$

$$(-a + (-b)i) + (a + bi) = (-a + a) + (-b + b)i = 0$$

Thus, each element of  $S$  has an additive inverse.

Commutativity. This follows from the fact that addition of complex numbers is commutative.

Multiplication makes  $S^\times$  into a group.

Associativity. This follows from the fact that multiplication of complex numbers is associative.

Identity. It is enough to notice that  $1 \in S^\times$ . Since  $1$  is the identity in  $\mathbb{C}^\times$ , it will also be the identity in  $S^\times$ .

Inverse. Let  $a + bi \in S^\times$ . Then  $a \neq 0$  or  $b \neq 0$ . Since  $a^2 + b^2$  is either 1, 2, 4, 5, 8,  $a^2 + b^2 \neq 0$  in  $\mathbb{F}_3$ . Thus, we can proceed as follows:

$$\frac{1}{a + bi} \cdot \frac{a - bi}{a - bi} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} \cdot i$$

Since  $a^2 + b^2 \neq 0$ , its inverse  $(a^2 + b^2)^{-1}$  in  $\mathbb{F}_3^\times$  exists. Therefore,  $a(a^2 + b^2)^{-1} \in \mathbb{F}_3$  and  $-b(a^2 + b^2)^{-1} \in \mathbb{F}_3$ . Thus,  $a(a^2 + b^2)^{-1} + (-b(a^2 + b^2)^{-1})i \in S$ . and the previous calculations show that this is the multiplicative inverse of  $a + bi$ .

Thus, each element of  $S$  has a multiplicative inverse.

Multiplication on  $S$  is commutative.

This follows from the fact that the multiplication of complex numbers is commutative.

Distributive Property.

This holds since it holds for the standard operations with complex numbers.

### $\mathbb{F}_5$ and $\mathbb{F}_7$ .

First notice that everything we did, excluding finding inverse, did not depend on the fact that  $p = 3$ . Thus, everything except that fact holds.

If  $p = 7$ , we can proceed the same way, because  $a^2 + b^2 \neq 0$  for all  $a, b \in \mathbb{F}_7$  such that at least one of them is not zero.

If  $p = 5$ , then this does not hold (take for example  $a = 2, b = 1$ ). Moreover, we can show that  $2 + i$  does not have an inverse!

Notice that

$$(2 + i)(2 - i) = 5 = 0$$

Now suppose that  $2 + i$  has the multiplicative inverse; denote it by  $(2 + i)^{-1}$ . Multiplying the above equality by it yields

$$2 - i = 0$$

This is a contradiction since clearly  $2 - i \neq 0$ . Thus,  $2 + i$  does not have the multiplicative inverse, so multiplication does not make this set into a group, so this set with standard complex addition and multiplication is not a field.

### **Result**

Denote the set of these numbers by  $S$ , and the set of its nonzero elements by  $S^\times$ . Check that:

$(S, +)$  is an abelian group.

$(S^\times, \cdot)$  is a group.

Multiplication is commutative.

The distributive property holds.

The same method works for  $\mathbb{F}_7$ , but it does not work with  $\mathbb{F}_5$ .

## Section 2

1. a

**(a)**

For any two scalars  $a, b$  and any vector  $v$ , we have that

$$(a + b)v = av + bv$$

Now take  $a = b = 0$ , the zero element of the field  $F$ . Then

$$(0 + 0)v = 0v + 0v$$

Now notice that  $0 + 0 = 0$ , since  $0$  is the zero element of the field. Thus,

$$0v = 0v + 0v$$

Since  $0v \in V$ , it has an inverse  $-0v$ . Add  $-0v$  to the above equality:

$$0v + (-0v) = 0v + 0v + (-0v)$$

Moreover,  $0v + (-0v) = 0$  (here it is the zero vector of  $V$ ). Thus,

$$0 = 0v + 0$$

Moreover,  $0v + 0 = 0v$ , so the above equality finally yields

$$\boxed{0v = 0}$$

Since  $v \in V$  was taken arbitrarily, the statement follows

**(b)**

For  $W$  to be a subspace, it must be closed under addition and scalar multiplication. So, since  $w \in W$ , we know that  $(-1)w \in W$ , since  $-1$  is a scalar. Moreover, since  $1w = w$ ,

$$w + (-1)w = 1w + (-1)w \stackrel{(1)}{=} (1 + (-1))w \stackrel{(2)}{=} 0w \stackrel{(a)}{=} 0$$

(1) follows from  $(a + b)v = av + bv$ .

(2) follows from  $1 + (-1) = 0$  in a field  $F$ .

Similarly,

$$(-1)w + w = (-1)w + 1w = (-1 + 1)w = 0w = 0$$

Therefore,

$$w + (-1)w = (-1)w + w = 0$$

This means that  $(-1)w$  is the inverse of  $w$  in  $V^+$ . Therefore, since the inverse is unique,  $-w = (-1)w$ , so  $-w \in W$ , since  $(-1)w$  is in  $W$ .

## Result

**(a)** Hint: the equality  $(0 + 0)v = 0v + 0v$  holds (why?).

**(b)** Show that  $(-1)w = -w$ .

2. a



(a)

We will show that this set is closed under addition and scalar multiplication.

Closed under addition.

Let  $A, B$  be two symmetric matrices. Let  $C = A + B$ . Then

$$C^t = (A + B)^t = A^t + B^t = A + B = C,$$

where the second equality follows from the properties of matrix transpose, while the third equality follows from the fact that  $A$  and  $B$  are symmetric.

Therefore,  $C = A + B$  is symmetric, so this set is closed under addition.

Closed under scalar multiplication.

Let  $A$  be some symmetric matrix, let  $c$  be some scalar. Let  $B = cA$ . Then

$$B^t = (cA)^t = cA^t = cA = B,$$

where the second equality follows from the properties of matrix transpose, while the third equality follows from the fact that  $A$  is symmetric.

Therefore,  $B = cA$  is symmetric, so this set is closed under scalar multiplication.

Conclusion.

The set of symmetric matrices truly is a subset of  $F^{n \times n}$ .

(b)

We can almost immediately see that this is not a subset of  $F^{n \times n}$ . Take any invertible matrix  $A$ . Multiply it by the zero element of  $F$ :  $0 \cdot A = 0$ , where  $0$  on the RHS is the zero matrix. However, the zero matrix is not invertible, so this set is not closed under scalar multiplication, so it is not a subspace of  $F^{n \times n}$ .

(c)

We will show that this set is closed under addition and scalar multiplication.

Closed under addition.

Let  $A, B$  be two upper triangular matrices. Let  $C = A + B$ . To show that  $C$  is upper triangular, we will show that for its coefficients  $c_{ij}$  the following holds:  $c_{ij} = 0$  whenever  $i > j$ .

By the definition of matrix addition,  $c_{ij} = a_{ij} + b_{ij}$ . Now, if  $i > j$ , then  $a_{ij} = b_{ij} = 0$  since  $A$  and  $B$  are upper triangular. So, when  $i > j$ , we conclude that  $c_{ij} = 0$ .

Therefore,  $C = A + B$  is upper triangular, so this set is closed under addition.

Closed under scalar multiplication.

Let  $A$  be some upper triangular matrix, let  $c$  be some scalar. Let  $B = cA$ . Then, the same as above,  $b_{ij} = c \cdot a_{ij}$  by the definition of the scalar multiplication. Furthermore, if  $i > j$ ,  $a_{ij} = 0$  since  $A$  is upper triangular, and  $b_{ij} = c \cdot a_{ij} = 0$ .

Therefore,  $B = cA$  is upper triangular, so this set is closed under scalar multiplication.

Conclusion.

The set of upper triangular matrices truly is a subset of  $F^{n \times n}$ .

## Result

(a) Yes.

(b) No.

(c) Yes.

## Section 3

1. a

Let  $A$  be some symmetric matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Transpose is:

$$A^t = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$$

Since  $A = A^t$ , we conclude that  $a_{ij} = a_{ji}$ , for  $i \neq j$  (for example,  $a_{12} = a_{21}$ ). Thus, we can write

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$$



Remember that  $a_{ij}$  are scalars, so we can simplify  $A$  even further:

$$A = a_{11} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \dots + a_{nn} \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} +$$

$$+ a_{12} \begin{bmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \dots + a_{1n} \begin{bmatrix} 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix} + \dots + a_{2n} \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 \end{bmatrix} + \dots$$

Now define  $E_i$  as a matrix with zero everywhere except at  $(i, i)$  where there is 1, and  $E_{ij}$  as a matrix with zero everywhere except at  $(i, j)$  and  $(j, i)$  where there is 1. Then

$$A = \sum_{i=1}^n a_{ii} E_i + \sum_{1 \leq i < j \leq n} a_{ij} E_{ij}$$

This means that the set

$$\{E_i \mid i \in \{1, 2, \dots, n\}\} \cup \{E_{ij} \mid 1 \leq i < j \leq n\}$$

spans the entire space of symmetric matrices (the fact that it is a subset of the set of symmetric matrices is clear from the definition of  $E_i$  and  $E_{ij}$ ).

Now we only need to see that this set is linearly independent. Take any linear relation

$$E_1 x_1 + \dots + E_n x_n + E_{12} x_{12} + \dots + E_{1n} x_{1n} + \dots + E_{2n} x_{2n} + \dots + E_{(n-1),n} x_{(n-1),n} = 0$$

for some elements  $x_i, x_{ij} \in \mathbb{R}$ , and 0 on the right side is a zero matrix. Using the definitions of matrices  $E_i$  and  $E_{ij}$ , the left side is

$$\begin{bmatrix} x_1 & x_{12} & \dots & x_{1n} \\ x_{12} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{nn} \end{bmatrix}$$

Since this must be equal to the zero matrix, we get  $x_i = 0$  for all  $i = 1, 2, \dots, n$ , and  $x_{ij} = 0$  for all  $1 \leq i < j \leq n$ . Thus, this set is linearly independent, so it is also a basis for the space of symmetric matrices.

## Result

3 of 3

Define  $E_i$  as a matrix with zero everywhere except at  $(i, i)$  where there is 1, and  $E_{ij}$  as a matrix with zero everywhere except at  $(i, j)$  and  $(j, i)$  where there is 1. Then the basis is

$$\{E_i \mid i \in \{1, 2, \dots, n\}\} \cup \{E_{ij} \mid 1 \leq i < j \leq n\}$$

2. a

Let

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_3 \end{bmatrix}$$

The equation  $AX = 0$  is equivalent to the system of linear equations

$$\begin{aligned} 2x_1 + x_2 + 2x_3 + 3x_4 &= 0 \\ x_1 + x_2 + 3x_3 &= 0 \end{aligned}$$

Subtract the second equation from the first:

$$\begin{aligned} x_1 &\quad - x_3 + 3x_4 = 0 \\ x_1 + x_2 + 3x_3 &= 0 \end{aligned}$$

Now subtract the first equation from the second:

$$\begin{aligned} x_1 &\quad - x_3 + 3x_4 = 0 \\ x_2 + 4x_3 - 3x_4 &= 0 \end{aligned}$$

Therefore,

$$x_1 = x_3 - 3x_4$$

and

$$x_2 = -4x_3 + 3x_4$$

This yields

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_3 - 3x_4 \\ -4x_3 + 3x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

Thus, every solution of  $AX = 0$  is of the above form. Since  $x_3$  and  $x_4$  are scalars, we conclude that the set

$$S = \left\{ \begin{bmatrix} 1 \\ -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

spans the set of all solutions of  $AX = 0$  (the fact that  $S$  is a subset of the space of all solutions of  $AX = 0$  is clear, since

$$A \begin{bmatrix} 1 \\ -4 \\ 1 \\ 0 \end{bmatrix} = 0 = A \begin{bmatrix} -3 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

To confirm that this is the basis of the set (space) of all solutions of  $AX = 0$ , we only need to confirm that it is linearly independent. To do this, we show that the linear relation

$$a \begin{bmatrix} 1 \\ -4 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

has only  $a = b = 0$  as a solution. Notice that this is equivalent to

$$\begin{bmatrix} a - 3b \\ -4a + 3b \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, from the third and the fourth component, we get that  $a = b = 0$ , as required. Thus, the set  $S$  is linearly independent, so it is also the basis for the set (space) of solutions  $AX = 0$ .

## Result

3 of 3

$$\left\{ \begin{bmatrix} 1 \\ -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

### 3. a

We need to show that the linear relation

$$ax^2 + b \cos x + ce^x = 0 \quad (1)$$

has only the trivial solution  $a = b = c$ .

First of all, since we are observing the set of functions, the equation (1) must hold for all real  $x$ . Specially, it must hold for  $x = 0$ ,  $x = \pm \frac{\pi}{2}$ . Picking such  $x$  yields the following equations:

$$\begin{aligned} x = 0 : & \quad b + c = 0 \\ x = \frac{\pi}{2} : & \quad a \cdot \frac{\pi^2}{4} + c \cdot e^{\pi/2} = 0 \\ x = -\frac{\pi}{2} : & \quad a \cdot \frac{\pi^2}{4} + c \cdot e^{-\pi/2} = 0 \end{aligned}$$

Subtracting the third equation from the second, we obtain the equation

$$c(e^{\pi/2} - e^{-\pi/2}) = 0$$

Since  $e^{\pi/2} - e^{-\pi/2} \neq 0$ , we conclude that  $c = 0$ . Plugging this into the second equation, we obtain

$$a \cdot \frac{\pi^2}{4} = 0 \implies a = 0$$

Plugging  $c = 0$  into the first equation we immediately obtain  $b = 0$ .

Thus, the only solution of the linear relation (1) is  $a = b = c = 0$ , so the given functions are linearly independent by definition.

## Result

Show that the linear relation

$$ax^2 + b \cos x + ce^x = 0$$

has only trivial solution  $a = b = c = 0$ . (Hint: the above equation must hold for all real  $x$ .)

4. a

We will first prove the result when  $A'$  is the result of applying **one** elementary row operation on  $A$ .

Let the rows of  $A$  be  $r_1, r_2, \dots, r_m$ . Let the rows of  $A'$  be  $r'_1, r'_2, \dots, r'_m$ . We have three cases regarding the elementary row operation used.

**1st case.** If we swapped the  $i$ th with the  $j$ th row, then  $r_k = r'_k$  for  $k \neq i, k \neq j$ , and  $r'_i = r_j, r'_j = r_i$ . However, this means that

$$\{r'_1, r'_2, \dots, r'_m\} = \{r_1, r_2, \dots, r_m\}$$

so

$$\text{span}\{r'_1, r'_2, \dots, r'_m\} = \text{span}\{r_1, r_2, \dots, r_m\}$$

is trivial.

**2nd case.** If we multiplied the  $i$ th row by a constant  $c \neq 0$ , then  $r'_k = r_k$  for  $k \neq i$ , and  $r'_i = cr_i$ . So,

$$\{r'_1, r'_2, \dots, r'_i, \dots, r'_m\} = \{r_1, r_2, \dots, cr_i, \dots, r_m\}$$

We first show that

$$\{r_1, r_2, \dots, cr_i, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_m\}$$

Let  $x \in \{r_1, r_2, \dots, cr_i, \dots, r_m\}$ . If  $x \neq cr_i$ , then  $x \in \{r_1, \dots, r_m\}$ , so  $x \in \text{span}\{r_1, \dots, r_m\}$  is trivial. If  $x = cr_i$ , then  $x \in \text{span}\{r_1, \dots, r_m\}$ , since  $cr_i$  is a linear combination of elements of  $\{r_1, \dots, r_m\}$ , so it is in subspace spanned by this set.

Thus,  $\{r_1, r_2, \dots, cr_i, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_m\}$ . Now we use Lemma 3.4.5 (note that  $\text{span}\{r_1, \dots, r_m\}$  is a subspace of  $\mathbb{R}^m$ ) to conclude that

$$\text{span}\{r_1, r_2, \dots, cr_i, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_m\}$$

On the other hand, we now show that

$$\{r_1, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, cr_i, \dots, r_m\}$$

Let  $x \in \{r_1, \dots, r_m\}$ . If  $x \neq r_i$ , then  $x \in \{r_1, \dots, cr_i, \dots, r_m\}$ , and  $x \in \text{span}\{r_1, \dots, cr_i, \dots, r_m\}$ . If  $x = r_i$ , then notice that, since  $c \neq 0$ ,  $r_i = \frac{1}{c} \cdot (cr_i)$ , so  $r_i \in \text{span}\{r_1, \dots, cr_i, \dots, r_m\}$ .

Thus,  $\{r_1, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, cr_i, \dots, r_m\}$ . By Lemma 3.4.5,

$$\text{span}\{r_1, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, cr_i, \dots, r_m\}$$

Therefore,

$$\text{span}\{r_1, \dots, r_m\} = \text{span}\{r_1, \dots, cr_i, \dots, r_m\},$$

which we needed to show.

**3rd case.** If we multiplied the  $j$ th row by  $c$  and added it to the  $i$ th row, then  $r'_k = r_k$  for  $k \neq i$ , and  $r'_i = r_i + cr_j$ . So,

$$\{r'_1, r'_2, \dots, r'_i, \dots, r'_m\} = \{r_1, r_2, \dots, r_i + cr_j, \dots, r_m\}$$

Now we use the methods used in the second case.

We first show that

$$\{r_1, r_2, \dots, r_i + cr_j, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_m\}$$

Let  $x \in \{r_1, r_2, \dots, r_i + cr_j, \dots, r_m\}$ . If  $x \neq r_i + cr_j$ , then  $x \in \{r_1, \dots, r_m\}$ , so  $x \in \text{span}\{r_1, \dots, r_m\}$  is trivial. If  $x = r_i + cr_j$ , then  $x \in \text{span}\{r_1, \dots, r_m\}$ , since  $r_i + cr_j$  is a linear combination of elements of  $\{r_1, \dots, r_m\}$ , so it is in subspace spanned by this set.

Thus,  $\{r_1, r_2, \dots, r_i + cr_j, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_m\}$ . Now we use Lemma 3.4.5 (note that  $\text{span}\{r_1, \dots, r_m\}$  is a subspace of  $\mathbb{R}^m$ ) to conclude that

$$\text{span}\{r_1, r_2, \dots, r_i + cr_j, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_m\}$$

On the other hand, we now show that

$$\{r_1, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_i + cr_j, \dots, r_m\}$$

Let  $x \in \{r_1, \dots, r_m\}$ . If  $x \neq r_i$ , then  $x \in \{r_1, \dots, r_i + cr_j, \dots, r_m\}$ , and  $x \in \text{span}\{r_1, \dots, r_i + cr_j, \dots, r_m\}$ . If  $x = r_i$ , then notice that  $r_i = (r_i + cr_j) + (-c)r_j$ , which is a linear combination of elements of  $\{r_1, \dots, r_i + cr_j, \dots, r_m\}$ , so  $r_i \in \text{span}\{r_1, \dots, r_i + cr_j, \dots, r_m\}$ .

Thus,  $\{r_1, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_i + cr_j, \dots, r_m\}$ . By Lemma 3.4.5,

$$\text{span}\{r_1, \dots, r_m\} \subseteq \text{span}\{r_1, \dots, r_i + cr_j, \dots, r_m\}$$

Therefore,

$$\text{span}\{r_1, \dots, r_m\} = \text{span}\{r_1, \dots, r_i + cr_j, \dots, r_m\},$$

which we needed to show.

### Multiple elementary row operations.

Now suppose that we had  $k$  elementary row operations; let  $A^{(i)}$  be the matrix after  $i$ th elementary row operations. Thus,  $A^{(k)} = A'$ . Abusing notation a bit for simplicity, let  $\text{rowspan } M$  denote the subspace spanned by rows of matrix  $M$ . Using the above result, we conclude that

$$\text{rowspan } A = \text{rowspan } A^{(1)}$$

and

$$\text{rowspan } A^{(i)} = \text{rowspan } A^{(i+1)}, i = 1, 2, \dots, k-1$$

Therefore,

$$\text{rowspan } A = \text{rowspan } A^{(k)} = \text{rowspan } A',$$

as required.

### Result

5 of 5

Hint: first show the result when only one elementary row operation was used. Then conclude the general result. (For example, by induction.)



Let  $(e_1, e_2, \dots, e_m)$  be a basis of  $W$  (such exists by Proposition 3.4.16 (a)). Now we can add elements to it to obtain a basis of  $\mathbb{R}^n$  (again, by Proposition 3.4.16 (a)); let  $(e_1, e_2, \dots, e_m, e_{m+1}, \dots, e_n)$  be some basis of  $\mathbb{R}^n$ .

Now take  $P$ , a basechanging matrix, which changes the standard basis to  $(e_1, \dots, e_n)$ .

Now let

$$B = \begin{bmatrix} 0_m & 0 \\ 0 & I_{n-m} \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix}$$

This is an  $n \times n$  matrix. Now we set  $A = BP$ . We will prove that the system

$$AX = 0$$

is the desired system.

Let  $Y \in W$ . Then  $Y = b_1 e_1 + \dots + b_m e_m$  for some scalars  $b_i$ , since  $(e_1, \dots, e_m)$  is a basis for  $W$ . By the definition of a basechanging matrix,

$$PY = \begin{bmatrix} b_1 \\ \vdots \\ b_m \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Now,

$$AY = BPY = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

Therefore,  $Y$  is a solution of  $AX = 0$ .

Now let  $X \in \mathbb{R}^n$  be a solution of  $AX = 0$ . Since  $X \in \mathbb{R}^n$ , and  $(e_1, \dots, e_n)$  is a basis of  $\mathbb{R}^n$ , there exist scalars  $a_i$  such that  $X = a_1e_1 + \dots + a_ne_n$ . By the definition of a basechanging matrix,

$$PX = \begin{bmatrix} b_1 \\ \vdots \\ b_m \\ b_{m+1} \\ \vdots \\ b_n \end{bmatrix}$$

Now,

$$0 = AX = BPX = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \\ b_{m+1} \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ b_{m+1} \\ \vdots \\ b_n \end{bmatrix}$$

So, we have  $b_{m+1} = \dots = b_n = 0$ , and

$$X = b_1e_1 + \dots + b_me_m \in W$$

This proves that  $X \in W$  if and only if  $AX = 0$ , which solves the problem.

## Result

3 of 3

Take basis of  $W$  and  $\mathbb{R}^n$  such that the basis of  $W$  is a subset of the basis of  $\mathbb{R}^n$ ; denote the basis of  $\mathbb{R}^n$  by  $S$ . Take the basechanging matrix  $P$  from the standard basis to the basis  $S$ . Consider  $A = BP$ , where

$$B = \begin{bmatrix} 0_m & 0 \\ 0 & I_{n-m} \end{bmatrix}$$

6. a



Let  $W$  be the space of solutions of this equation. Let

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

be in  $W$ . Then

$$x_1 + 2x_2 + \dots + nx_n = 0 \implies x_1 = -2x_2 - \dots - nx_n$$

Therefore,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + x_n \begin{bmatrix} -n \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (1)$$

Now let

$$S = \left( \begin{bmatrix} -2 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} -n \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right)$$

Every element of  $S$  solves the given equation, and by (1) it spans the entire space of solutions. To conclude that this is a basis of  $W$ , we need to check that it is linearly independent.

Let

$$a_2 \begin{bmatrix} -2 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_n \begin{bmatrix} -n \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2)$$

be the linear relation. We need to prove that  $a_2 = \dots = a_n = 0$ .

Since

$$a_2 \begin{bmatrix} -2 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_n \begin{bmatrix} -n \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} -2a_2 - \dots - na_n \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

the equation (2) becomes

$$\begin{bmatrix} -2a_2 - \dots - na_n \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

From this, we easily confirm that  $a_2 = \dots = a_n = 0$ , so the set  $S$  truly is linearly independent, and a basis of  $W$ .

## Result

$$S = \left( \begin{bmatrix} -2 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} -n \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right)$$

7. a

Since the set

$$S = (X_1 Y_1^t, \dots, X_1 Y_n^t, \dots, X_m Y_n^t)$$

consists of  $mn$  vectors, and  $\dim \mathbb{R}^{m \times n} = mn$ , by Proposition 3.4.21 (b) we only need to see that  $S$  spans the entire  $\mathbb{R}^{m \times n}$ .

Now let  $A \in \mathbb{R}^{m \times n}$ . Notice that, for fixed  $i = 1, 2, \dots, m$ ,

$$A^t X_i$$

is well-defined since  $A^t$  has  $m$  columns, and  $X_i$  is a  $m \times 1$  matrix, and  $A^t X_i$  is a  $n \times 1$  matrix. Therefore, we can say that  $A^t X_i \in \mathbb{R}^n$ . Since  $(Y_1, Y_2, \dots, Y_n)$  is a basis for  $\mathbb{R}^n$ , so it also spans  $\mathbb{R}^n$ , there exist scalars  $c_{ij}$ ,  $j = 1, \dots, n$ , such that

$$A^t X_i = c_{i1} Y_1 + \dots + c_{in} Y_n$$

Using the properties of a transpose of a matrix,

$$X_i^t A = X_i^t (A^t)^t = (A^t X_i)^t = (c_{i1} Y_1 + \dots + c_{in} Y_n)^t = c_{i1} Y_1^t + \dots + c_{in} Y_n^t$$

Now multiply this equation by  $X_j$  from the left and use the distributivity of matrix multiplication:

$$X_j X_i^t A = c_{i1} X_j Y_1^t + \dots + c_{in} X_j Y_n^t$$

Therefore, for each  $i, j \in \{1, 2, \dots, m\}$  we have that

$$X_j X_i^t A \in \text{span } S$$

Now notice that

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} [0 \ 0 \ \dots \ 1 \ \dots \ 0] = \begin{bmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

where 1 on the left side is in the  $i$ th component, and on position  $(i, i)$  on the right side. If we set

$$E_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

and

$$E_{ii} = \begin{bmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix},$$

we see that

$$I_m = E_{11} + \dots + E_{mm} = E_1 E_1^t + \dots + E_m E_m^t$$

Moreover,  $E_i \in \mathbb{R}^m$ , and, since  $(X_1, \dots, X_m)$  is a basis of  $\mathbb{R}^m$ , it also spans  $\mathbb{R}^m$ , so there exist scalars  $a_{i1}, \dots, a_{im}$  such that

$$E_i = a_{i1} X_1 + \dots + a_{im} X_m$$

Thus,

$$\begin{aligned} E_i E_i^t &= (a_{i1} X_1 + \dots + a_{im} X_m)(a_{i1} X_1^t + \dots + a_{im} X_m^t) \\ &= a_{i1}^2 X_1 X_1^t + \dots + a_{im} a_{i1} X_m X_1^t + \dots + a_{im}^2 X_m X_m^t \\ &= \sum_{j=1}^m \sum_{k=1}^m a_{ij} a_{ik} X_j X_k \end{aligned}$$

Finally,

$$\begin{aligned} A &= I_m A \\ &= \left( \sum_{i=1}^m E_{ii} \right) A \\ &= \sum_{i=1}^m E_{ii} A \\ &= \sum_{i=1}^m E_i E_i^t A \\ &= \sum_{i=1}^m \left( \sum_{j=1}^m \sum_{k=1}^m a_{ij} a_{ik} X_j X_k \right) A \\ &= \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ij} a_{ik} X_j X_k A \end{aligned}$$

Since  $X_j X_k A \in \text{span } S$ , we conclude that also  $\sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ij} a_{ik} X_j X_k A \in \text{span } S$ , so  $A \in \text{span } S$ . This means that  $S$  spans the entire  $\mathbb{R}^{m \times n}$ , and, by the discussion from the start, it also means that  $S$  is a basis of  $\mathbb{R}^{m \times n}$ .

## Result

This is truly a basis for  $\mathbb{R}^{m \times n}$ . Hint: Proposition 3.4.21 (b).

8. a

Since the set  $(v_1, \dots, v_n)$  has  $n$  vectors, and  $\dim \mathbb{R}^n = n$ , by Proposition 3.4.21 it follows that  $(v_1, \dots, v_n)$  is a basis if and only if  $(v_1, \dots, v_n)$  is linearly independent.

Let  $v_i = (a_{i1}, \dots, a_{in})$ . Then

$$A = [v_1 \ v_2 \ \dots \ v_n] = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$$

Now we see that

$$A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c_1 v_1 + \dots + c_n v_n \quad (1)$$

(this is seen by direct computation). This system and its equivalent form will prove useful.

### *A is invertible.*

Suppose that  $A$  is invertible; that is,  $A^{-1}$  exists. To prove that  $(v_1, \dots, v_n)$  is linearly independent, take any linear relation

$$c_1 v_1 + \dots + c_n v_n = 0 \quad (2)$$

We must prove that  $c_1 = \dots = c_n = 0$ . By (1), the equation (2) is equivalent to

$$A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = 0$$

Since  $A^{-1}$  exists, we can multiply the above equation by it:

$$A^{-1} A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = A^{-1}(0)$$

So,

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = 0$$

which prove that  $c_1 = \dots = c_n = 0$ , and the set  $(v_1, \dots, v_n)$  is invertible.

$(v_1, \dots, v_n)$  is linearly independent.

Suppose that  $(v_1, \dots, v_n)$  is linearly independent. This means that the equation

$$c_1 v_1 + \dots + c_n v_n = 0$$

has only the trivial solution  $c_1 = \dots = c_n = 0$ . This also means, by (I), that

$$A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = 0$$

has only the trivial solution

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = 0$$

. By Theorem 1.2.21, we now conclude that  $A$  is invertible, as required.

### Result

Hint: show that

$$A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c_1 v_1 + \dots + c_n v_n$$

## Section 4

1. a

(a)

Since  $\mathbf{B}$  has 3 vectors, and  $\dim \mathbb{R}^3 = 3$ , by Proposition 3.4.21 (c) we only need to confirm that  $\mathbf{B}$  is linearly independent.

So, we need to show that the linear relation

$$a \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + b \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + c \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has only the trivial solution  $a = b = c = 0$ . Since this relation is equivalent to

$$\begin{bmatrix} a + 2b + 3c \\ 2a + b + c \\ 2b + c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

we obtain the system of equations

$$\begin{aligned} a + 2b + 3c &= 0 \\ 2a + b + c &= 0 \\ 2b + c &= 0 \end{aligned}$$

Multiply the first equation by  $-2$  and add it to the second:

$$\begin{aligned} a + 2b + 3c &= 0 \\ -3b - 5c &= 0 \\ 2b + c &= 0 \end{aligned}$$

Multiply the third equation by 5 and add it to the second:

$$\begin{aligned} a + 2b + 3c &= 0 \\ 7b &= 0 \\ 2b + c &= 0 \end{aligned}$$

Now from the second equation we get  $b = 0$ . Plugging this into the third equation we obtain  $c = 0$ . Finally, from the first equation we now get  $a = 0$ . Thus,

$$\boxed{a = b = c = 0}$$

and  $\mathbf{B}$  is linearly independent. Therefore, it is also a basis for  $\mathbb{R}^3$ .

(b)

The coordinate vector of  $v$  with respect to  $\mathbf{B}$  is

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

where  $x_1, x_2, x_3$  are such that

$$v = x_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

This equation is equivalent to

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 2x_1 + x_2 + x_3 \\ 2x_2 + x_3 \end{bmatrix}$$

From this, we obtain the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 2 \\ 2x_2 + x_3 &= 3 \end{aligned}$$

Multiply the first equation by  $-2$  and add it to the second:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ -3x_2 - 5x_3 &= 0 \\ 2x_2 + x_3 &= 3 \end{aligned}$$

Multiply the third equation by 5 and add it to the second:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 7x_2 &= 15 \\ 2x_2 + x_3 &= 3 \end{aligned}$$

From the second equation, we obtain  $x_2 = \frac{15}{7}$ . Plugging this into the third equation,

$$\frac{30}{7} + x_3 = 3 \implies x_3 = -\frac{9}{7}$$

Plugging all this into the first equation,

$$x_1 + \frac{30}{7} - \frac{27}{7} = 1 \implies x_1 = \frac{4}{7}$$

Therefore, the coordinate vector of  $v$  is

$$X = \begin{bmatrix} 4/7 \\ 15/7 \\ -9/7 \end{bmatrix}$$



(c)

By definition, we must find a matrix  $P$  such that

$$\mathbf{B}' = \mathbf{B}P$$

So,  $P$  is such that (write the vectors of corresponding bases as columns of corresponding matrices):

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix} P$$

From Exercise 3.3.8 we know that  $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix}$  is invertible. Thus,

$$P = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Now we find the inverse. We use the following [elementary row operations](#):

1. Interchange the  $i$ th and the  $j$ th rows:  $R_i \leftrightarrow R_j$ .
2. Multiply the  $i$ th row by a scalar  $c \neq 0$ :  $cR_i$ .
3. Multiply the  $j$ th row by a scalar  $c$  and add it to the  $i$ th row:  $R_i + cR_j$ .

$$\begin{aligned} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] &\xrightarrow{-2R_1+R_2} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -5 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \\ &\xrightarrow{5R_3+R_2} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 7 & 0 & -2 & 1 & 5 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \\ &\xrightarrow{\frac{1}{7}R_2} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{7} & \frac{1}{7} & \frac{5}{7} \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \\ &\xrightarrow{-2R_2+R_3} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{7} & \frac{1}{7} & \frac{5}{7} \\ 0 & 0 & 1 & \frac{4}{7} & -\frac{2}{7} & -\frac{3}{7} \end{array} \right] \\ &\xrightarrow{-2R_2+R_1} \left[ \begin{array}{ccc|ccc} 1 & 0 & 3 & \frac{11}{7} & -\frac{2}{7} & -\frac{10}{7} \\ 0 & 1 & 0 & -\frac{2}{7} & \frac{1}{7} & \frac{5}{7} \\ 0 & 0 & 1 & \frac{4}{7} & -\frac{2}{7} & -\frac{3}{7} \end{array} \right] \\ &\xrightarrow{-3R_3+R_1} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{7} & \frac{4}{7} & -\frac{1}{7} \\ 0 & 1 & 0 & -\frac{2}{7} & \frac{1}{7} & \frac{5}{7} \\ 0 & 0 & 1 & \frac{4}{7} & -\frac{2}{7} & -\frac{3}{7} \end{array} \right] \end{aligned}$$

Therefore,

$$P = \begin{bmatrix} -\frac{1}{7} & \frac{4}{7} & -\frac{1}{7} \\ -\frac{2}{7} & \frac{1}{7} & \frac{5}{7} \\ \frac{4}{7} & -\frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{7} & -\frac{2}{7} & \frac{3}{7} \\ \frac{1}{7} & \frac{3}{7} & -\frac{3}{7} \\ -\frac{2}{7} & \frac{1}{7} & \frac{5}{7} \end{bmatrix}$$

## Result

(a) Check that  $\mathbf{B}$  is linearly independent. Why is it enough to conclude that it is also a basis for  $\mathbb{R}^3$ ?

(b)

$$X = \begin{bmatrix} 4/7 \\ 15/7 \\ -9/7 \end{bmatrix}$$

(c)

$$P = \begin{bmatrix} 4/7 & -2/7 & 2/7 \\ 1/7 & 3/7 & -3/7 \\ -2/7 & 1/7 & 6/7 \end{bmatrix}$$

2. a

(a)

First of all,

$$e_1 + e_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$e_1 - e_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

To solve the exercise, we must find a matrix  $P$  such that

$$\mathbf{B} = \mathbf{E}P$$

Writing the vectors of corresponding basis as columns of a corresponding matrix, we get the equation

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} P = I_2 P$$

From this we immediately get

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

(b)

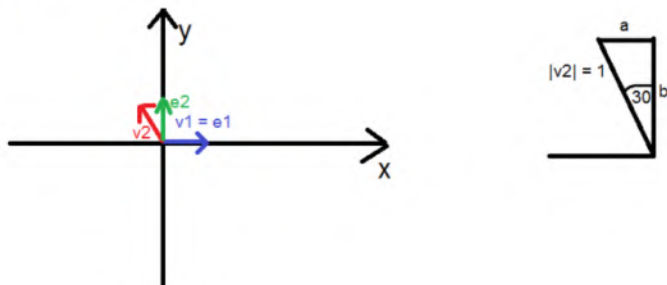
Similarly to (a), we obtain the equation

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} P = I_n P$$

From this we immediately get

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

This vector  $v_2$  is equal to  $v_2 = (-1/2, \sqrt{3}/2)^t$ . This is obtained using the following sketch (all that is left to notice is that  $v_2 = ae_1 + be_2$  and  $a = \sin 30^\circ$ ,  $b = \cos 30^\circ$ ):



As before, we obtain the equation

$$\begin{bmatrix} 1 & -1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} P = I_2 P$$

From this we immediately get

$$P = \begin{bmatrix} 1 & -1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}$$

## Result

(a)

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

(b)

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

(c)

$$P = \begin{bmatrix} 1 & -1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}$$

3. a

Let  $\mathbf{B}'$  be the new basis,  $\mathbf{B}$  be the old basis. Let  $P$  be the *basechanging* matrix:

$$\mathbf{B}' = \mathbf{B}P$$

By Exercise 3.3.8, matrices  $B$  and  $B'$ , which are the matrices whose columns are vectors of corresponding basis, are invertible. Moreover, now clearly

$$B' = BP$$

and

$$P = B^{-1}B'$$

Now we conclude that  $P$  is invertible as a product of two invertible matrices.

This also means that now there exist elementary matrices  $E_1, \dots, E_m$  such that

$$PE_1 \cdots E_m = I_n$$

Therefore,

$$P = E_m^{-1} \cdots E_1^{-1}$$

Note additionally that all  $E_i^{-1}$  are also elementary matrices.

Finally, this means that

$$\mathbf{B}' = \mathbf{B}P = \mathbf{B}E_m^{-1} \cdots E_1^{-1}$$

All that is left to notice is that each  $E_i^{-1}$  affects  $\mathbf{B}$  in one of the described ways (this is because of the definition of elementary matrices), and that there is a finite number of them.

## Result

2

Hint: let  $\mathbf{B}'$  be the new basis, and  $P$  be the basechanging matrix:

$$\mathbf{B}' = \mathbf{B}P$$

Write  $P$  as a product of elementary matrices (why can you do that?). Why does this solve the problem?

4. a

(a)

Let  $\mathbf{B}$  be some fixed basis for  $\mathbb{F}_p^2$ . Define the function

$$f : GL_2(\mathbb{F}_p) \rightarrow \{\text{bases of } \mathbb{F}_p^2\}$$

as

$$f(P) = \mathbf{B}P$$

We will prove that such  $f$  is a bijection.

$f$  is surjective.

Let  $\mathbf{B}'$  be some other basis. Then there exists a basechanging matrix  $P$  such that

$$\mathbf{B}' = \mathbf{B}P$$

Now let  $B'$  and  $B$  be matrices whose columns are vectors of corresponding basis. Then, by the above equality,

$$B' = BP$$

Furthermore, by Exercise 3.3.8,  $B$  and  $B'$  are invertible, so

$$P = B^{-1}B'$$

is also invertible. Thus,  $P \in GL_2(\mathbb{F}_p)$ , since  $\det P \neq 0$  if and only if  $P$  is invertible. Hence,

$$f(P) = \mathbf{B}P$$

and  $f$  is a surjection.

(NOTE: You cannot just write  $f(P) = \mathbf{B}P$  before checking that  $P \in GL_2(\mathbb{F}_p)$ ; this is because we must make sure that  $P$  is in the domain of  $f$ !)

$f$  is injective.

Suppose that  $P_1, P_2 \in GL_2(\mathbb{F}_p)$  are such that  $f(P_1) = f(P_2) = \mathbf{B}'$ . Using the same notation as in the proof that  $f$  is surjective, we now obtain that

$$B' = BP_1$$

$$B' = BP_2$$

Thus,

$$P_1 = B^{-1}B' = P_2$$

Hence,  $P_1 = P_2$ , which means that  $f$  is injective.

Conclusion.

We now know that  $f$  is a bijection, which means that its domain and codomain have the same number of elements, which proves the statement of the exercise.

(b)

$GL_2(\mathbb{F}_p)$

If the columns are linearly dependent, they do not form a basis, hence the matrix is not invertible by Exercise 3.3.8; moreover, the matrix is also not in  $GL_2(\mathbb{F}_p)$ . Thus, if we denote columns by  $v_1$  and  $v_2$ , the linear relation

$$av_1 + bv_2 = 0 \quad (1)$$

must have only the trivial solution  $a = b = 0$ .

If  $v_1 = 0$ , then the linear relation (1) has a nontrivial solution  $a = 1$  and  $b = 0$ ; thus, the matrix will not be in  $GL_2(\mathbb{F}_p)$ . This means that  $v_1 \neq 0$ , and we have  $p^2 - 1$  combinations to accomplish this (since  $v_1 \in \mathbb{F}_p^2$ , in every component we can have  $p$  elements, so the total number of combinations is  $p^2$ ; however, we must remove the zero vector, making the number of legal combinations  $p^2 - 1$ ).

Now onto  $v_2$ . Notice that if  $v_2 = \lambda v_1$ , for some scalar  $\lambda$ , the linear relation (1) has a nontrivial solution  $a = -\lambda$ ,  $b = 1$  (since  $-\lambda v_1 + v_2 = 0$ ).

On the other hand, suppose that  $xv_1 + yv_2 = 0$  for some nontrivial  $x, y$  (at least one of them is nonzero). Then we must have  $y \neq 0$ ; if  $y = 0$ , then  $x \neq 0$ , and  $xv_1 = 0$  implies  $v_1 = 0$ , which is a contradiction. Thus, we can write

$$v_2 = -xy^{-1}v_1$$

or

$$v_2 = \lambda v_1$$

Therefore, the linear relation (1) has a nontrivial solution if and only if  $v_2 = \lambda v_1$ , for some scalar  $\lambda \in \mathbb{F}_p$ . Since we have  $p$  such scalars,  $p$  vectors are not allowed to be  $v_2$ . Since there are  $p^2$  total combinations to pick  $v_2$ , the number of legal combinations is  $p^2 - p$ .

Finally, the total number of combinations of both  $v_1$  and  $v_2$  are

$$(p^2 - 1)(p^2 - p) = (p - 1)(p + 1)p(p - 1) = p(p + 1)(p - 1)^2$$

$SL_2(\mathbb{F}_p)$

Let

$$M = \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix}$$

, where  $i \in \{2, 3, \dots, p - 1\}$ . Define a function

$$f : SL_2(\mathbb{F}_p) \rightarrow \{A \in GL_2(\mathbb{F}_p) \mid \det A = i\}$$

by

$$f(A) = AM$$

*is well-defined.*

We first check that  $f$  is well-defined. However,

$$\det f(A) = \det A \det \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot i = i,$$

since  $A \in SL_2(\mathbb{F}_p)$ . Thus, it is really well-defined.

*is injective.*

Let  $A, B \in SL_2(\mathbb{F}_p)$  be such that

$$f(A) = f(B)$$

This means that

$$AM = BM$$

Since  $\det M \neq 0$ , it is invertible; multiply the above equation by  $M^{-1}$  from the right:

$$A = B$$

This proves that  $f$  is injective.

$f$  is surjective.

Let  $X$  be a matrix such that  $\det X = i$ . Consider

$$XM^{-1}$$

Since  $\det(AB) = \det A \det B$ , and  $MM^{-1} = I$ , we conclude that

$$\det(MM^{-1}) = \det I \implies \det M \det M^{-1} = 1 \implies i \det M^{-1} = 1 \implies \det M^{-1} = i^{-1}$$

Therefore,

$$\det(XM^{-1}) = \det X \det M^{-1} = i \cdot i^{-1} = 1$$

So,  $XM^{-1} \in SL_2(\mathbb{F}_p)$ . Furthermore,

$$f(XM^{-1}) = XM^{-1}M = X$$

Thus,  $f$  is surjective.

### Conclusion.

Since  $f$  is bijective, the number of matrices with the determinant 1 is the same as the number of matrices with the determinant  $i$ . This means that the set of matrices with the determinant  $i$  has the same number of elements as  $SL_2(\mathbb{F}_p)$ . Notice that  $i \in \{2, 3, \dots, p-1\}$  was arbitrarily taken; we will need it later.

Now write

$$GL_2(\mathbb{F}_p) = SL_2(\mathbb{F}_p) \cup \{A \in GL_2(\mathbb{F}_p) \mid \det A = 2\} \cup \dots \cup \{A \in GL_2(\mathbb{F}_p) \mid \det A = p-1\} \quad (*)$$

and notice that the sets on the right side are pair-wise disjoint.

Denote by  $\text{card } X$  the number of elements of  $X$ . Moreover, notice that  $\text{card}(X \cup Y) = \text{card } X + \text{card } Y$  when  $X$  and  $Y$  are disjoint. This leads to:

$$\text{card}(X_1 \cup \dots \cup X_n) = \text{card } X_1 + \dots + \text{card } X_n$$

when  $X_i$  are pair-wise disjoint.



Finally, if we denote  $D_i = \{A \in GL_2(\mathbb{F}_p) \mid \det A = i\}$ , from the equality (\*) we conclude that

$$\text{card } GL_2(\mathbb{F}_p) = \text{card } SL_2(\mathbb{F}_p) + \text{card } D_2 + \dots + \text{card } D_{p-1} \quad (**)$$

From before,

$$\text{card } GL_2(\mathbb{F}_p) = p(p+1)(p-1)^2$$

Moreover,  $\text{card } D_i = \text{card } SL_2(\mathbb{F}_p)$  for  $i = 2, \dots, p-1$ . Thus, (\*\*) becomes

$$p(p+1)(p-1)^2 = (p-1)\text{card } SL_2(\mathbb{F}_p)$$

Finally,

$$\text{card } SL_2(\mathbb{F}_p) = p(p+1)(p-1),$$

as required.

## Result

6 of 6

(a) Use the fact that we can "change" from one basis to another by using basechanging matrices.

(b) For  $GL_2(\mathbb{F}_p)$ , use Exercise 3.3.8. For  $SL_2(\mathbb{F}_p)$ , prove that the number of matrices of determinant 1 is the same as the number of matrices of determinant  $i$ , for  $i = 2, \dots, p-1$ .

5. a

(a)

Dimension = 0.

There is only one subspace of dimension 0:  $\{0\}$ .

Dimension = 1.

The subspaces  $W$  of dimension 1 are of the form  $W = \text{span}\{v_1\}$ , where  $v_1 \neq 0$  (because the dimension is the number of elements of a basis for  $W$ ). Since

$$\text{span}\{v_1\} = \{\alpha v_1 \mid \alpha \in \mathbb{F}_p\}$$

( $\alpha \in \mathbb{F}_p$  because  $\mathbb{F}_p$  is the field in this exercise), we conclude that

$$W = \{0, v_1, \dots, (p-1)v_1\}$$

Now notice that

$$W = \text{span}\{av_1\},$$

for any  $a = 1, 2, \dots, p-1$ . This is because  $av_1 \in \text{span}\{v_1\}$ , so  $\{av_1\} \subseteq \text{span}\{v_1\}$ . Moreover, now  $\text{span}\{av_1\} \subseteq \text{span}\{v_1\}$  and, because the dimensions are the same,  $\text{span}\{av_1\} = \text{span}\{v_1\}$ . Therefore,  $W$  is uniquely determined by picking any of its nonzero points.

Now to conclude something about the number of said  $W$ . There are  $p^3 - 1$  nonzero vectors  $v \in \mathbb{F}_p^3$ . It is easy to see that

$$\{W \mid W \text{ is a subspace of dimension 1}\} = \{\text{span}\{v\} \mid v \neq 0\}$$

However,  $p - 1$  vectors  $v$  span the same subspace, so there are

$$\frac{p^3 - 1}{p - 1} = \frac{(p^2 + p + 1)(p - 1)}{p - 1} = p^2 + p + 1$$

such subspaces.

Dimension = 2.

Similarly to the dimension 1, all subspaces of dimension 2 are of the form

$$W = \text{span}\{v, w\},$$

where  $v$  and  $w$  are linearly independent. There are  $(p^3 - 1)$  choices for  $v \neq 0$ , and  $(p^3 - p)$  choices for  $w$ . This is because  $v, w$  are linearly dependent on if and only if  $w = \lambda v$ , for some  $\lambda \in \mathbb{F}_p$ . To prove this, first notice that if  $w = \lambda v$ , then  $-\lambda v + w = 0$ ; thus, the linear relation  $av + bw = 0$  has nontrivial solution. On the other hand, if  $av + bw = 0$  has a nontrivial solution, then  $b = 0$  implies  $av = 0$ , which, because  $v \neq 0$ , implies  $a = 0$ , so we get a trivial solution. So, we must have that  $b \neq 0$ . Then  $w = -\frac{a}{b}v$ , so  $w = \lambda v$ . This proves that  $\{v, w\}$ , with  $v \neq 0$ , is linearly dependent if and only if  $w = \lambda v$ , so there are  $p$  such  $w$  (since  $\lambda \in \mathbb{F}_p$ ).

Moreover, in this discussion the order in which  $v$  and  $w$  are taken mattered; when counting subspaces  $W$ , the order in which the elements are in the basis does not matter, so the number drops to

$$\frac{(p^3 - 1)(p^3 - p)}{2} = \frac{p(p - 1)^2(p^2 + p + 1)(p + 1)}{2}$$

However, since we get the same result when taking  $av$  and  $bw$ , with  $a \neq 0, b \neq 0$  (because the span is the same), the number drops to

$$\frac{\frac{p(p-1)^2(p^2+p+1)(p+1)}{2}}{(p-1)^2} = \frac{p(p^2 + p + 1)(p + 1)}{2}$$

(because there are  $p - 1$  such  $a \neq 0$ , and  $p - 1$  such  $b \neq 0$ ).

The question is how many  $\{v, w\}$  span the same subspace  $W$  of dimension 2. Notice that  $\text{span}\{v\}$  and  $\text{span}\{w\}$  are two subspaces of dimension 1 in  $W$ . Similarly as in dimension 3, we obtain that there are  $\frac{p^2 - 1}{p - 1} = p + 1$  different subspaces of dimension 1. So, we see that each basis of  $W$  gets us two different subspaces of dimension 1 of  $W$ . The converse is also true: by choosing two dimensions 1 of  $W$ , by making the union of their bases, we get a basis for  $W$  (because they must be linearly dependent).

When picking two subspaces out of them  $p + 1$ , we get

$$\binom{p + 1}{2} = \frac{(p + 1)p}{2}$$

different combinations.

So, the total number of subspaces  $W$  of dimension 2 is

$$\frac{\frac{p(p^2 + p + 1)(p + 1)}{2}}{\frac{(p + 1)p}{2}} = p^2 + p + 1$$

Dimension = 3.

Only  $\mathbb{F}_p^3$  is a subspace of  $\mathbb{F}_p^3$  of dimension 3.

(b)

Dimension = 0.

There is only one subspace of dimension 0:  $\{0\}$ .

Dimension = 1.

As in (a), there are

$$\frac{p^4 - 1}{p - 1} = p^3 + p^2 + p + 1$$

such subspaces (the only difference is that the total number of nonzero vectors is  $p^4 - 1$ , not  $p^3 - 1$ ).

Dimension = 2.

This is also very similar to (a). There are

$$\frac{(p^4 - 1)(p^4 - p)}{2} = \frac{p(p - 1)^2(p^3 + p^2 + p + 1)(p^2 + p + 1)}{2}$$

"starting"  $v$  and  $w$ , using

$$p^4 - p = p(p^3 - 1) = p(p - 1)(p^2 + p + 1)$$

. We must divide this number by  $(p - 1)^2$ . Thus, the number drops to

$$\frac{p(p^3 + p^2 + p + 1)(p^2 + p + 1)}{2}$$

Again, there are

$$\binom{p + 1}{2} = \frac{p(p + 1)}{2}$$

pairs of different subspaces of dimension 1. Thus, the total number is

$$\frac{\frac{p(p^3 + p^2 + p + 1)(p^2 + p + 1)}{2}}{\frac{p(p + 1)}{2}} = \frac{(p^3 + p^2 + p + 1)(p^2 + p + 1)}{p + 1} = (p^2 + 1)(p^2 + p + 1),$$

using

$$\frac{p^3 + p^2 + p + 1}{p + 1} = \frac{(p^2 + 1)(p + 1)}{p + 1} = p^2 + 1$$

### Dimension = 3.

This is very similar to dimension 2, but we have three vectors  $\{v, w, z\}$ . We can pick  $v, w$  in  $(p^4 - 1)(p^4 - p)$ . If  $z = \alpha v + \beta w$ , then the linear relation has a nontrivial solution. On the other hand, if the linear relation  $av + bw + cz = 0$  has a nontrivial solution, we  $c = 0$  implies  $av + bw = 0$ , so  $a = b = 0$  since  $v, w$  are linearly independent. Thus,  $c \neq 0$ , and  $z = \alpha v + \beta w$ . Since there are  $p^2$  choices on  $\alpha, \beta$  ( $p$  on each of them), we get that there are  $p^4 - p^2$  independent  $z$ 's. Therefore, the "starting" number is (do not forget to divide by  $3! = 6$ , since the order does not matter)

$$\frac{(p^4 - 1)(p^4 - p)(p^4 - p^2)}{6} = \frac{p^3(p - 1)^3(p^3 + p^2 + p + 1)(p^2 + p + 1)(p + 1)}{6}$$

Since we can scale  $v, w, z$  with nonzero scalars and obtain the same subspace, we divide this number by  $(p - 1)^3$ :

$$\frac{\frac{p^3(p-1)^3(p^3+p^2+p+1)(p^2+p+1)(p+1)}{6}}{(p-1)^3} = \frac{p^3(p^3 + p^2 + p + 1)(p^2 + p + 1)(p + 1)}{6}$$

Like before, we now use that we have

$$p^2 + p + 1$$

subspaces of dimension 1 in  $W$  (this is shown in (a); we can somewhat take  $W$  as if we have  $\mathbb{F}_p^3$ ). We must pick 3 subspaces of dimension 1 in  $W$  which are different and the third must not be in the span of the first two. So, we first take two of them:

$$\binom{p^2 + p + 1}{2} = \frac{(p^2 + p + 1)(p^2 + p)}{2}$$

Now there are  $p^2 + p + 1$  total subspaces of dimension 1; however,  $p + 1$  of them will be in the span of the first two (there are  $p + 1$  subspaces of dimension 1 in a subspace of dimension 2). Thus, the number becomes

$$\frac{(p^2 + p + 1)(p^2 + p)}{2} \cdot (p^2 + p + 1 - (p + 1)) = \frac{p^3(p^2 + p + 1)(p + 1)}{2}$$

However, we must further divide this number by 3 (it does not matter which of the three vectors will span the subspace of dimension 1):

$$\frac{p^3(p^2 + p + 1)(p + 1)}{6}$$

Finally, the total number of subspaces of dimension 3 is

$$\frac{\frac{p^3(p^3 + p^2 + p + 1)(p^2 + p + 1)(p + 1)}{6}}{\frac{p^3(p^2 + p + 1)(p + 1)}{6}} = p^3 + p^2 + p + 1$$

### Dimension = 4.

There is only one subspace of dimension 4:  $\mathbb{F}_p^4$ .

## Result

(a) Dimension 0: 1

Dimension 1:  $p^2 + p + 1$

Dimension 2:  $p^2 + p + 1$

Dimension 3: 1

(b) Dimension 0: 1

Dimension 1:  $p^3 + p^2 + p + 1$

Dimension 2:  $p^2 + 1$

Dimension 3:  $p^3 + p^2 + p + 1$

Dimension 4: 1

## Section 5

1. a

!!!

2. a

We first find a basis for  $W_1$ .

Let  $A \in W_1$ ,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Since the trace of  $A$  is zero, we get

$$a_{11} + \dots + a_{nn} = 0 \implies a_{11} = -a_{22} - \dots - a_{nn}$$

Thus,

$$A = a_{12} \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + a_{22} \begin{bmatrix} -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \dots + a_{nn} \begin{bmatrix} -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Thus, if we denote by  $E_{ij}$  the matrix which has 1 on position  $(i, j)$  and 0 everywhere else, we get

$$A = \sum_{i \neq j} a_{ij} E_{ij} + \sum_{i=2}^n a_{ii} (E_{ii} - E_{11})$$



Since each  $E_{ij}$ ,  $i \neq j$  and  $E_{ii} - E_{11}$  is clearly in  $W_1$ , we conclude that the set

$$S = \{E_{ij} \mid i, j \in \{1, \dots, n\}, i \neq j\} \cup \{E_{ii} - E_{11} \mid i \in \{1, \dots, n\}\}$$

spans  $W_1$ . Moreover, notice that

$$\sum_{i \neq j} a_{ij} E_{ij} + \sum_{i=1}^n (E_{ii} - E_{11}) = \begin{bmatrix} -\sum_{i=2}^n a_{ii} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Thus, the linear relation

$$\sum_{i \neq j} a_{ij} E_{ij} + \sum_{i=1}^n (E_{ii} - E_{11}) = 0$$

is equivalent to

$$\begin{bmatrix} -\sum_{i=2}^n a_{ii} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = 0$$

from which  $a_{ij} = 0$  easily follows. Thus, the set  $S$  is linearly independent, so it is also a basis for  $W_1$ .

Moreover, notice that there are  $n^2 - 1$  elements in  $S$ . Now we define

$$T = S \cup \{E_{11}\}$$

Then  $T$  has  $n^2$  elements. We will prove that it is a basis for  $\mathbb{R}^{n \times n}$ . For this, it is sufficient to prove that it is linearly independent (because  $\dim \mathbb{R}^{n \times n} = n^2$ ). As before, the linear relation

$$\sum_{i \neq j} a_{ij} E_{ij} + a_{11} E_{11} + \sum_{i=1}^n (E_{ii} - E_{11}) = 0$$

is equivalent to

$$\begin{bmatrix} a_{11} - \sum_{i=2}^n a_{ii} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = 0$$

From this we easily get  $a_{ij} = 0$ , for all  $i, j \in \{1, 2, \dots, n\}$ . Thus,  $T$  is linearly independent, and a basis of  $\mathbb{R}^{n \times n}$ .

Now, by Proposition 3.6.4. (a), we can set

$$W_2 = \text{span}\{E_{11}\} = \left\{ \begin{bmatrix} a & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}$$

## Result

$$W_2 = \left\{ \begin{bmatrix} a & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}$$

We only need to show that  $W_1, \dots, W_k$  are independent, that is

$$w_1 + \dots + w_k = 0$$

for  $w_i \in W_i, i = 1, \dots, k$ , implies that  $w_1 = w_2 = \dots = w_k = 0$ .

Notice that

$$w_k = (-w_1) + \dots + (-w_{k-1}) \quad (1)$$

Therefore,  $w_k \in W_1 + \dots + W_{k-1}$  (since  $-w_i \in W_i$ ). Therefore,

$$w_k \in W_k \cap (W_1 + \dots + W_{k-1})$$

However,  $W_k \cap (W_1 + \dots + W_{k-1}) = \{0\}$ , so  $w_k = 0$ . From (1) we now get (after multiplying by  $-1$ , of course)

$$w_1 + \dots + w_{k-1} = 0$$

Inductively, we show that

$$w_1 = \dots = w_k = 0,$$

as required.

## Result

Let

$$w_1 + \dots + w_k = 0,$$

for  $w_i \in W_i$ . From this, show that  $w_k = (W_1 + \dots + W_{k-1}) \cap W_k$ , so  $w_k = 0$ . Complete the proof.

# Section 6

## 1. a

The span consists of vectors of the form

$$v = c_1 v_1 + \dots + c_n v_n, \quad (1)$$

where  $n$  is some positive integer,  $c_1, \dots, c_n \in \mathbb{R}$ , and  $v_1, \dots, v_n \in (w, e_1, e_2, \dots)$ .

Now let  $v \in \text{span}(w, e_1, e_2, \dots)$ . Then, by the above definition,

$$v = cw + c_1 e_{i_1} + \dots + c_n e_{i_n}, \quad (2)$$

where  $c, c_1, c_2, \dots, c_n \in \mathbb{R}$ ,  $i_1, \dots, i_n$  are positive integers. (If  $w$  does not originally appear in (1), we can write it in (2) as  $0 \cdot w$ .)

Notice that  $v$  has  $c_j + c$  on  $i_j$ th coordinate, and  $c$  everywhere else. So, we now define the set  $S$  as a set of all vectors  $x \in \mathbb{R}^\infty$  such that there exists some positive integer  $k$  such that for all positive integers  $m \geq k$  we have that  $m$ th coordinate of  $x$  is equal to  $d$ , for some  $d \in \mathbb{R}$  (that is,  $x$  is a constant sequence after some point). Notice that  $v \in S$  (with  $m = \max\{i_1, \dots, i_n\} + 1$ ). Therefore, since  $v$  was arbitrarily taken,

$$\text{span}(w, e_1, e_2, \dots) \subseteq S$$



We want to prove that  $S \subseteq \text{span}(w, e_1, e_2, \dots)$  also holds. Let  $x \in S$ . Denote by  $x(i)$  the  $i$ th coordinate of  $x$ . Let  $c \in \mathbb{R}$  and the positive integer  $k$  be such that  $m \geq k$  implies  $x(m) = c$ , for all positive integers  $m$  (such  $k, c$  exists by definition of  $S$ ). Now all that is left to notice is that

$$x = (x(1) - c)e_1 + \dots + (x(k-1) - c)e_{k-1} + cw$$

Truly, let  $n$  be some positive integer  $n$ . If  $n < k$ , by definitions of  $e_i$  and  $w$ , we conclude that  $x(n) = (x(n) - c) + c = x(n)$ , because only  $e_n$  and  $w$  have 1 on the  $n$ th coordinate, and all other  $e_i$  have zeros on this coordinate. If  $n \geq k$ , then  $x(n) = c$ , because  $w$  has 1 on the  $n$ th coordinate, and all  $e_i$  have zeros. Thus,  $x \in \text{span}(w, e_1, e_2, \dots)$ , and

$$S \subseteq \text{span}(w, e_1, e_2, \dots)$$

This means that

$$\text{span}(w, e_1, e_2, \dots) = S$$

## Result

This span is the set of all sequences which are constant after some point.

## 2. a

### Definition of a function.

Denote by  $V$  the vector space of doubly infinite row vectors. By definition of an isomorphism, we must find a bijective function  $\varphi : V \rightarrow \mathbb{R}^\infty$  such that

$$\varphi(v + w) = \varphi(v) + \varphi(w) \quad \text{and} \quad \varphi(cv) = c\varphi(v) \quad (1)$$

for all vectors  $v, w \in V$  and scalars  $c \in \mathbb{R}$ .

We will define

$$\varphi(\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots) = (a_0, a_{-1}, a_1, a_{-2}, a_2, \dots)$$

To clarify, let  $k \in \mathbb{Z}$ . If  $k < 0$ , then we send the  $k$ th coordinate of  $(\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots)$  to the  $2|k|$ th coordinate, and if  $k \geq 0$ , we send it to the  $(2k+1)$ th coordinate of the vector of  $\mathbb{R}^\infty$ . We will now prove that  $\varphi$  is a bijection and that conditions (1) hold.

### $\varphi$ injective?

Let  $(a), (b) \in V$  be such that  $\varphi(a) = \varphi(b)$ ; that is,

$$(a_0, a_{-1}, a_1, a_{-2}, a_2, \dots) = (b_0, b_{-1}, b_1, b_{-2}, b_2, \dots)$$

From this we easily get that  $a_k = b_k$  for all  $k \in \mathbb{Z}$ . This also means that

$$(a) = (\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots) = (\dots, b_{-2}, b_{-1}, b_0, b_1, b_2, \dots) = (b)$$

Therefore,  $(a) = (b)$ , so  $\varphi$  is injective.

### $\varphi$ surjective?

Let  $(y) \in \mathbb{R}^\infty$ ;  $(y) = (y_1, y_2, \dots)$ . We define

$$(x) = (\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots) = (\dots, y_4, y_2, y_1, y_3, y_5, \dots)$$

To clarify, let  $k \in \mathbb{Z}$ . If  $k < 0$ , then the  $k$ th coordinate of  $(x)$  is the  $(2|k|)$ th coordinate of  $(y)$ . If  $k \geq 0$ , then the  $k$ th coordinate of  $(x)$  is the  $(2k+1)$ th coordinate of  $(y)$ .

Now we easily check that

$$\varphi(x) = (y),$$

so  $\varphi$  is surjective.

$\varphi$  satisfies the conditions (1)?

Let  $(v), (w) \in V$ . Then

$$\begin{aligned}\varphi(v+w) &= \varphi(\dots, v_{-2}+w_{-2}, v_{-1}+w_{-1}, v_0+w_0, v_1+w_1, v_2+w_2, \dots) \\ &= (v_0+w_0, v_{-1}+w_{-1}, v_1+w_1, v_{-2}+w_{-2}, v_2+w_2, \dots) \\ &= (v_0, v_{-1}, v_1, v_{-2}, v_2, \dots) + (w_0, w_{-1}, w_1, w_{-2}, w_2, \dots) \\ &= \varphi(v) + \varphi(w)\end{aligned}$$

So,

$$\varphi(v+w) = \varphi(v) + \varphi(w)$$

for all vectors  $v, w \in V$ .

Now let  $(v) \in V, c \in \mathbb{R}$ . Then

$$\begin{aligned}\varphi(cv) &= \varphi(\dots, cv_{-2}, cv_{-1}, cv_0, cv_1, cv_2, \dots) \\ &= (cv_0, cv_{-1}, cv_1, cv_{-2}, cv_2, \dots) \\ &= c(v_0, v_{-1}, v_1, v_{-2}, v_2, \dots) \\ &= c\varphi(v)\end{aligned}$$

So,

$$\varphi(cv) = c\varphi(v)$$

for all  $v \in V$  and  $c \in \mathbb{R}$ .

### Conclusion.

The function  $\varphi$  which we defined is truly an isomorphism between  $V$  and  $\mathbb{R}^\infty$ , so  $V$  is isomorphic to  $\mathbb{R}^\infty$ .

### Result

Hint: show that

$$\varphi(\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots) = (a_0, a_{-1}, a_1, a_{-2}, a_2, \dots)$$

is an isomorphism.

### 3. a

To prove that  $l^p$  is a proper subspace of  $l^{p+1}$ , we need to prove that:

1.  $l^p$  is a subset of  $l^{p+1}$ .
2.  $l^p$  is closed under the addition of vectors.
3.  $l^p$  is closed under the scalar multiplication.
4. There exists some vector in  $l^{p+1}$  which is not in  $l^p$ .

### 1. Subset?

Let  $(a) = (a_1, a_2, \dots) \in l^p$ . Since the series

$$\sum_{n=1}^{\infty} |a_n|^p$$

converges, we must have that

$$\lim_{n \rightarrow \infty} |a_n|^p = 0,$$

and, because of that,

$$\lim_{n \rightarrow \infty} |a_n| = 0$$

Because of that, there exists some  $N \in \mathbb{N}$  such that, for  $n \in \mathbb{N}$ ,

$$(n \geq N) \text{ implies } (|a_n| < 1)$$

(pick  $\varepsilon = 1$  in the definition of the limit). Now write

$$\sum_{n=1}^{\infty} |a_n|^{p+1} = \sum_{n=1}^N |a_n|^{p+1} + \sum_{n=N+1}^{\infty} |a_n|^{p+1}$$

However,  $|a_n| < 1$  for  $n \geq N$ , so  $|a_n|^{p+1} < |a_n|^p$  for  $n \geq N$ . Therefore,

$$\sum_{n=1}^N |a_n|^{p+1} + \sum_{n=N+1}^{\infty} |a_n|^{p+1} < \sum_{n=1}^N |a_n|^{p+1} + \sum_{n=N+1}^{\infty} |a_n|^p$$

Furthermore, because  $|a_n| \geq 0$  for all  $n \in \mathbb{N}$ ,

$$\sum_{n=N+1}^{\infty} |a_n|^p \leq \sum_{n=1}^{\infty} |a_n|^p < \infty$$

Now we conclude that

$$\sum_{n=1}^N |a_n|^{p+1} + \sum_{n=N+1}^{\infty} |a_n|^p < \infty$$

and

$$\sum_{n=1}^N |a_n|^{p+1} + \sum_{n=N+1}^{\infty} |a_n|^{p+1} < \sum_{n=1}^N |a_n|^{p+1} + \sum_{n=N+1}^{\infty} |a_n|^p < \infty$$

Finally,

$$\sum_{n=1}^{\infty} |a_n|^{p+1} < \infty,$$

so  $(a) \in l^{p+1}$ , as required.

So,  $l^p \subseteq l^{p+1}$  holds.

## 2. Closed under addition?

Let  $(a), (b) \in l^p$ . Then

$$\begin{aligned} |a_n + b_n|^p &\leq (|a_n| + |b_n|)^p \\ &\leq (2 \cdot \max\{|a_n|, |b_n|\})^p \\ &= 2^p \cdot \max\{|a_n|^p, |b_n|^p\} \\ &\leq 2^p(|a_n|^p + |b_n|^p) \end{aligned}$$

The first inequality is the Triangle Inequality, the second inequality follows from  $|a_n| \leq \max\{|a_n|, |b_n|\}$  and  $|b_n| \leq \max\{|a_n|, |b_n|\}$ , the final inequality holds because  $|a_n|^p \geq 0$  and  $|b_n|^p \geq 0$ , so  $|a_n|^p \leq |a_n|^p + |b_n|^p$  and  $|b_n|^p \leq |a_n|^p + |b_n|^p$ , which implies that  $\max\{|a_n|^p, |b_n|^p\} \leq |a_n|^p + |b_n|^p$ . Since

$$\sum_{n=1}^{\infty} |a_n|^p < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} |b_n|^p < \infty,$$

and  $|a_n + b_n|^p \leq 2^p(|a_n|^p + |b_n|^p)$ , we conclude that

$$\sum_{n=1}^{\infty} |a_n + b_n|^p \leq \sum_{n=1}^{\infty} (2^p(|a_n|^p + |b_n|^p)) < \infty$$

Therefore,  $(a) + (b) \in l^p$ , as required.

## 3. Closed under scalar multiplication?

Let  $(a) \in l^p$ ,  $c \in \mathbb{R}$ . Then

$$\sum_{n=1}^{\infty} |ca_n|^p = |c|^p \sum_{n=1}^{\infty} |a_n|^p < \infty,$$

because  $\sum_{n=1}^{\infty} |a_n|^p < \infty$ . Thus,  $c \cdot (a) \in l^p$ .

## 4. Proper subset?

Let  $(a)$  be defined with  $a_n = \frac{1}{n^{1/p}}$ . Then

$$\sum_{n=1}^{\infty} |a_n|^{p+1} = \sum_{n=1}^{\infty} \frac{1}{n^{(p+1)/p}} < \infty,$$

since  $\frac{p+1}{p} > 1$ .

On the other hand,

$$\sum_{n=1}^{\infty} |a_n|^p = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

Thus,  $(a) \in l^{p+1} \setminus l^p$ .

(Note that in this exercise we used the following Theorem: for  $\alpha \in \mathbb{R}$ ,

$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$  converges if and only if  $\alpha > 1$ .)

## Result

To prove that  $l^p$  is a proper subspace of  $l^{p+1}$ , prove the following:

1.  $l^p$  is a subset of  $l^{p+1}$ .
2.  $l^p$  is closed under the addition of vectors.
3.  $l^p$  is closed under the scalar multiplication.
4. There exists some vector in  $l^{p+1}$  which is not in  $l^p$ .

4. a

Let  $S$  be countably infinite such that

$$V = \text{span } S$$

Suppose that  $U$  is an uncountably infinite subset of  $V$  which is linearly independent. Then we can extend it to the basis  $\mathbf{B}$  for  $V$ ;  $\mathbf{B}$  must then also be uncountably infinite, since  $U \subseteq \mathbf{B}$ .

Since  $\mathbf{B}$  is a basis for  $V$ , it must span the entire  $V$ . Specially, we can write every  $s \in S$  as a linear combination of vectors of  $\mathbf{B}$ :

$$s = c_1 v_1 + \dots + c_n v_n,$$

for some  $n \in \mathbb{N}$ ,  $c_1, \dots, c_n \in F$ ,  $v_1, \dots, v_n \in \mathbf{B}$ .

Now define  $T$  as a set of all vectors of  $\mathbf{B}$  which are present in some linear combination

$$c_1 v_1 + \dots + c_n v_n \in S,$$

for some  $n \in \mathbb{N}$ ,  $c_1, \dots, c_n \in F$ ,  $v_1, \dots, v_n \in \mathbf{B}$ .

(So, we write each  $s \in S$  as a linear combination of vectors of  $\mathbf{B}$  and "put" vectors used in that linear combination into  $T$ .)

First note that  $T \subseteq \mathbf{B}$ , by definition.

Since  $S$  is countably infinite and every linear combination has a finite number of vectors, we conclude that  $T$  is also countably infinite. Furthermore, by definition of  $T$ ,

$$S \subseteq \text{span } T$$

Since  $\text{span } S = V$ , we also conclude that  $\text{span } T = V$ . Therefore,  $T$  spans the entire  $V$ .

Since  $T$  is countably infinite and  $\mathbf{B}$  is uncountably infinite, we obtain that  $T \subset \mathbf{B}$  (proper subset). By the above, we now obtained that there exists a proper subset of  $\mathbf{B}$  which spans the entire  $V$ , which is absurd since  $\mathbf{B}$  is a basis.

So, we obtained a contradiction, so the assumed  $U$  cannot exist. This means that all linearly independent subsets of  $V$  are either finite or countably infinite.

## Result

2 of 2

Hint: suppose that  $U$  is an uncountably infinite linearly independent subset and extend it to a basis  $\mathbf{B}$  for  $V$ . Try to obtain a contradiction (for example, find a proper subset of  $\mathbf{B}$  which spans the entire  $V$ ).

# Miscellaneous Problem

1. a

The function  $\det$  is surjective.

To prove that  $\det$  is surjective, we need to show that for each  $i \in \mathbb{F}_p$ , there exists some matrix  $A$  such that  $\det A = i$ .

Let  $i \in \mathbb{F}_p$ , and let

$$A = \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix}$$

Then

$$\det A = i,$$

as required.

All nonzero values are obtained the same number of times.

Let  $i \in \{2, \dots, p-1\}$ . Define a function

$$f : \{\text{matrices of determinant } 1\} \rightarrow \{\text{matrices of determinant } i\}$$

by

$$f(M) = ME_i,$$

where

$$E_i = \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix}$$

. First of all, we check that  $f$  is well-defined; that is, that  $\det(ME_i) = i$ . However, this is clear by the properties of the determinant:

$$\det(ME_i) = \det M \det E_i = 1 \cdot i = i,$$

since  $\det M = 1$ .

Now we want to see that  $f$  is a bijection. To prove that, we first prove that it is injective.

Let  $M_1, M_2$  be such that  $f(M_1) = f(M_2)$ ; that is,

$$M_1E_i = M_2E_i$$

Since  $\det E_i \neq 0$ , it is invertible; so, we can multiply the above equality by  $E_i^{-1}$  from the right to get

$$M_1 = M_2,$$

which proves that  $f$  is injective.



To prove that  $f$  is surjective, let  $N$  be such that  $\det N = i$ . Then we consider the matrix

$$NE_i^{-1}$$

First of all, by the properties of the determinant:

$$E_i E_i^{-1} = I_2 \implies \det(E_i E_i^{-1}) = \det I_2 \implies \det E_i \det E_i^{-1} = 1 \implies i \det E_i^{-1} = 1$$

Therefore,  $\det E_i^{-1} = i^{-1}$ . This means that

$$\det(NE_i^{-1}) = \det N \det E_i^{-1} = ii^{-1} = 1$$

So, we can conclude that

$$f(NE_i^{-1}) = NE_i^{-1} E_i = N$$

and that  $f$  is surjective. (NOTE: we cannot immediately conclude that  $f(NE_i^{-1}) = N$  without first checking that  $NE_i^{-1}$  is in the domain of  $f$ !)

Since  $f$  is a bijection, its domain and codomain have the same number of elements. Since  $i$  was taken arbitrarily, we now conclude that each nonzero number is truly taken the same number of times.

#### There are more matrices of determinant 0 than of determinant 1.

We will first find the number of all matrices with nonzero determinant.

Let  $A$  be such matrix. Its first column must not be equal to the zero column; thus, we have  $p^2 - 1$  options for the first column.

Now onto the second column. We now that  $\det A = 0$  if and only if  $A$  is invertible, which, by Exercise 3.3.8, is if and only if its columns are the basis of  $\mathbb{F}_p^2$ , which is clearly true if and only if the columns are linearly independent.

Now let  $v, w$  be the columns of  $A$ , with  $v$  being the first column. Thus,  $v \neq 0$ . If  $w = \lambda v$ , for some scalar  $\lambda$ , then they are clearly not linearly independent, since  $(-\lambda)v + 1 \cdot w = 0$  and  $1 \neq 0$ .

Suppose that  $v$  and  $w$  are linearly dependent; that is, there exist some scalars  $a, b$ , not both zero, such that

$$av + bw = 0$$

If  $b = 0$ , then  $av = 0$ , which means that  $a = 0$  since  $v \neq 0$ . This is a contradiction. Thus,  $b \neq 0$ , and  $w = -\frac{a}{b}v = \lambda v$ .

So, we conclude that  $v$  and  $w$  are linearly dependent if and only if  $w = \lambda v$  for some  $\lambda \in \mathbb{F}_p$  (because  $\lambda$  is a scalar).

Since there are  $p$  choices for  $\lambda$ , we have  $p$  columns  $w$  which are linearly dependent with  $v$ . Thus, the number of linearly independent columns is  $p^2 - p$ .

From all this, we conclude that the number of matrices with the nonzero determinant is equal to

$$(p^2 - 1)(p^2 - p) = p(p + 1)(p - 1)^2$$

Since all nonzero numbers are taken the same number of times, and there are  $p - 1$  nonzero numbers in  $\mathbb{F}_p$ , we get that we have

$$\frac{p(p + 1)(p - 1)^2}{p - 1} = p(p + 1)(p - 1) = p^3 - p$$

matrices with the determinant 1.



Total number of matrices in  $\mathbb{F}_p^{2 \times 2}$  is  $p^4$ . Of them,  $(p^2 - 1)(p^2 - p)$  are of nonzero determinant. This means that we have

$$p^4 - (p^2 - 1)(p^2 - p) = p^3 + p^2 - p$$

matrices with the determinant zero. Now it is clear that

$$p^3 + p^2 - p > p^3 - p,$$

so there are more matrices of determinant 0 than of determinant 1.

## Result

5 of 5

For the first part, notice that

$$\det \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix} = i$$

For the second part, prove that the following function is a bijection:

$$f : \{\text{matrices of determinant 1}\} \rightarrow \{\text{matrices of determinant } i\}$$

with

$$f(M) = ME_i$$

For the final part, prove that the number of matrices of determinant 1 is  $p^3 - p$ , while the number of matrices of determinant 0 is  $p^3 + p^2 - p$ .

## 2. a

Let  $I_n$  be the identity  $n \times n$  matrix. Define the set

$$S = \{I_n, A, A^2, \dots, A^{n^2}\}$$

We will show that  $S$  is linearly dependent.

Suppose that  $S$  is linearly independent. Then we can extend it to the basis of  $\mathbb{R}^{n \times n}$ ; let  $\mathbf{B}$  be such basis. Since  $S \subseteq \mathbf{B}$ ,  $\mathbf{B}$  has at least  $n^2 + 1$  elements. However,  $\dim \mathbb{R}^{n \times n} = n^2$ . Thus,  $\mathbf{B}$  cannot be a basis of  $\mathbb{R}^{n \times n}$ , which means that  $S$  cannot be linearly independent.

Now, we know that the linear relation

$$c_0 I_n + c_1 A + \dots + c_{n^2} A^{n^2} = 0 \quad (1)$$

has at least one nontrivial solution with regards to  $c_i$ ; fix one of them.

Let  $c_N$  be the nonzero  $c_i$  of the largest index (such exists because not all  $c_i$  are zero). Then (1) becomes

$$c_0 I_n + c_1 A + \dots + c_N A^N = 0$$

Since  $c_N \neq 0$ , then we can divide the above equality by it, and get

$$A^N + \dots + \frac{c_1}{c_N} A + \frac{c_0}{c_N} = 0,$$

as required.

### Result

Hint: show that the set

$$S = \{I_n, A, A^2, \dots, A^{n^2}\}$$

is not linearly independent.

### 3. a

#### (a)

We start with a little digression. Let  $\mathcal{P}_4$  be the set of all real polynomials (in variable  $t$ ) which are of degree 4 or less. Thus,

$$\mathcal{P}_4 = \{at^4 + bt^3 + ct^2 + dt + e \mid a, b, c, d, e \in \mathbb{R}\}$$

It is easy to show that this is a vector space, and we will show that

$$\mathbf{B} = (t^4, t^3, t^2, t, 1)$$

is a basis for  $\mathcal{P}_4$ .

First of all,

$$(at^4 + bt^3 + ct^2 + dt + e) \in \text{span } \mathbf{B},$$

so  $\mathbf{B}$  spans the entire  $\mathcal{P}_4$ . Now we want to show that it is linearly independent.

Take any linear relation

$$at^4 + bt^3 + ct^2 + dt + e \cdot 1 = 0$$

Then, by the Theorem of Equality of Polynomials, we get that

$$a = b = c = d = e = 0$$

This proves that  $\mathbf{B}$  is linearly independent.

Thus,  $\mathbf{B}$  is a basis for  $\mathcal{P}_4$ , and since  $\mathbf{B}$  has 5 elements we conclude that  $\dim \mathcal{P}_4 = 5$ .

Now onto the exercise. We define the set

$$S = \{x(t)^2, x(t)y(t), y(t)^2, x(t), y(t), 1\}$$

Notice that all elements of  $S$  are polynomials of degree 4 or less. Moreover,  $S$  must be linearly dependent!

Suppose that  $S$  is linearly independent, then we can extend it to the basis  $T \supseteq S$  of  $\mathcal{P}_4$ . However,  $T$  has at least 6 elements (because  $S$  has 6 elements), which is impossible since  $\dim \mathcal{P}_4 = 5$  (all bases must have exactly 5 elements). Therefore,  $S$  is linearly dependent.

Since  $S$  is not linearly independent, there exists a nontrivial linear relation

$$ax(t)^2 + bx(t)y(t) + cy(t)^2 + dx(t) + ey(t) + f = 0$$

Therefore,

$$f(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$$

is a nonzero polynomial such that

$$f(x(t), y(t)) = 0$$

(b)

Here we first conclude that

$$t^2 = x(t) + 1 \implies t = \sqrt{x(t) + 1}$$

(taking only the positive root is enough). Then,

$$y(t) = t(t^2 - 1) = \sqrt{x(t) + 1}x(t)$$

Squaring,

$$y(t)^2 = x(t)^3 + x(t)^2$$

So, we define

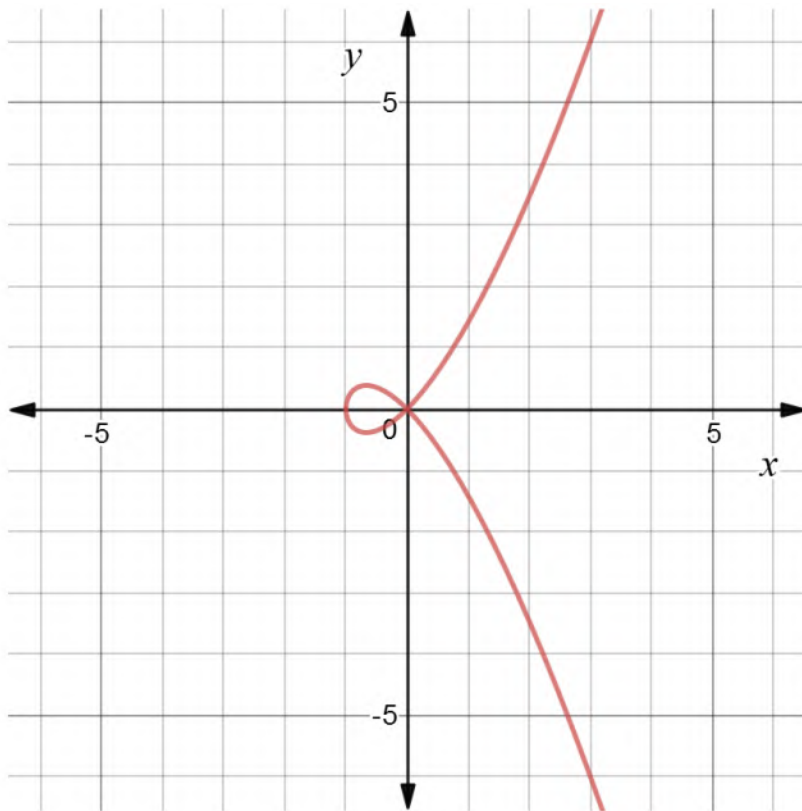
$$f(x, y) = x^3 + x^2 - y^2$$

and conclude that

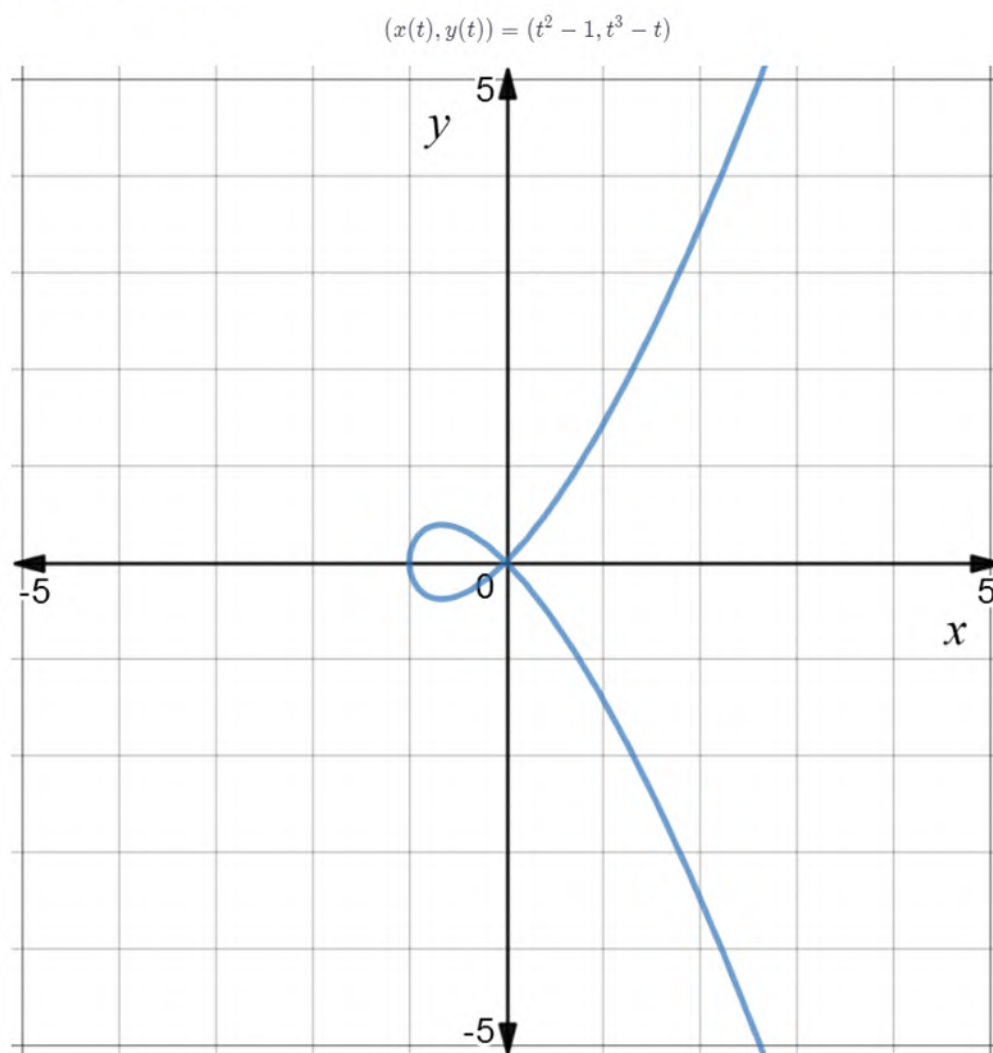
$$f(x(t), y(t)) = 0$$

Now we first sketch

$$\{f(x, y) = 0\} = \{x^3 + x^2 - y^2 = 0\}$$



Now we sketch the path



(c)

This is almost the same as (a). Let  $x(t)$  be of degree  $m$ ,  $y(t)$  be of degree  $n$ . We can prove that  $\mathcal{P}_{mn}$  of all real polynomials of degree  $mn$  or less is a vector space, and that  $\dim \mathcal{P}_{mn} = mn + 1$ . Now we define a set

$$S = \{x(t)^i y(t)^j \mid 0 \leq i \leq n, 0 \leq j \leq m\}$$

Notice that  $S$  has  $(m+1)(n+1) = mn + m + n + 1 > mn + 1$  elements. Thus,  $S$  cannot be linearly independent, so there exists some nontrivial linear relation

$$\sum_{i=0}^n \sum_{j=0}^m a_{ij} x(t)^i y(t)^j = 0$$

So, we define

$$f(x, y) = \sum_{i=0}^n \sum_{j=0}^m a_{ij} x^i y^j$$

(a) Prove that  $\mathcal{P}_4$  of real polynomials of degree 4 or less is a vector space and that its dimension is 5. Define  $S = \{x(t)^2, x(t)y(t), y(t)^2, x(t), y(t), 1\}$ . How can you, using the information provided, find a polynomial  $f$ ?

(b) For example,  $f(x, y) = x^3 + x^2 - y^2$ .

(c) HINT: Very similar to (a).

4. a

Suppose that

$$V = V_1 \cup V_2 \cup \dots \cup V_n,$$

where  $V_i$  are proper subspaces of  $V$ .

Let  $x \in V_1$ . Since  $V_1$  is a proper subspace of  $V$ , there exists some  $y \in V \setminus V_1$ . Now we consider the vectors  $x + ay$ , where  $a \in F$ , and  $F$  is the field of  $V$ . Since  $F$  is infinite, there are infinitely many vectors of the form  $x + ay$ . Moreover, if  $x + ay \in V_1$ , for  $a \neq 0$ , then  $(x + ay) - x \in V_1$ , since  $V_1$  is a subspace of  $V$ , so it must be closed under vector addition and scalar multiplication. This also means that  $ay \in V_1$ . Since  $V_1$  is closed under scalar multiplication, we conclude that  $y = \frac{1}{a}(ay) \in V_1$ , which is a contradiction. So, for all  $a \neq 0$  we conclude that  $x + ay \in V_2 \cup \dots \cup V_n$ .

This means that there exists some  $V_j$ ,  $j \in \{2, \dots, n\}$  such that at least two vectors of the form  $x + ay$  are in it. Let  $x + ay$  and  $x + by$  be in  $V_j$ , with  $a \neq b$ . Then

$$(x + by) - (x + ay) = (b - a)y \in V_j$$

Moreover,  $b - a \neq 0$ , so

$$y = \frac{1}{b - a}((b - a)y) \in V_j$$

Now we conclude that  $ay \in V_j$ , and

$$x = (x + ay) - (ay) \in V_j$$

Thus,  $x \in V_2 \cup \dots \cup V_n$ .

Therefore,

$$V_1 \subseteq V_2 \cup \dots \cup V_n$$

This means that

$$V = V_1 \cup V_2 \cup \dots \cup V_n = V_2 \cup \dots \cup V_n$$

Proceeding inductively, we can show that

$$V_i \subseteq \bigcup_{j=i+1}^n V_j,$$

so

$$V = V_n$$

This is impossible since  $V_n$  is a proper subset of  $V$ !

Therefore, we obtained a contradiction, so we cannot have that  $V$  is a finite union of its proper subspaces.

Suppose that

$$V = V_1 \cup V_2 \cup \dots \cup V_n$$

Let  $x \in V_1$ . By observing the vectors  $x + ay$ , where  $y \in V \setminus V_1$ , and  $a$  is a scalar, conclude that  $x \in V_2 \cup \dots \cup V_n$ . Thus,

$$V_1 \subseteq V_2 \cup \dots \cup V_n$$

and

$$V = V_1 \cup V_2 \cup \dots \cup V_n = V_2 \cup \dots \cup V_n$$

Complete the proof.

5. a

(a)

Suppose that  $c \neq 0$ . Then

$$x^3 - 2 = (cx^2 + bx + a) \left( \frac{1}{c}x - \frac{b}{c^2} \right) + \left( \frac{b^2}{c^2} - \frac{a}{c} \right) x + \left( \frac{ab}{c^2} - 2 \right)$$

Plug in  $x = \alpha$ . Then

$$\alpha^3 - 2 = (c\alpha^2 + b\alpha + a) \left( \frac{1}{c}\alpha - \frac{b}{c^2} \right) + \left( \frac{b^2}{c^2} - \frac{a}{c} \right) \alpha + \left( \frac{ab}{c^2} - 2 \right)$$

However,  $\alpha^3 - 2 = 0$  and  $(c\alpha^2 + b\alpha + a) = 0$  by assumption on  $a, b, c$ . Then

$$\left( \frac{b^2}{c^2} - \frac{a}{c} \right) \alpha + \left( \frac{ab}{c^2} - 2 \right) = 0 \quad (1)$$

If  $\left( \frac{b^2}{c^2} - \frac{a}{c} \right) \neq 0$ , then

$$\alpha = \frac{\frac{ab}{c^2} - 2}{\frac{b^2}{c^2} - \frac{a}{c}}$$

Since all number on the right side are rational, we conclude that  $\alpha$  is also rational. However, this is a contradiction since  $\alpha \notin \mathbb{Q}$ .

Therefore, we have that

$$\frac{b^2}{c^2} - \frac{a}{c} = 0 \implies ac = b^2 \implies a = \frac{b^2}{c}$$

and, plugging this into (1),

$$\frac{ab}{c^2} - 2 = 0 \implies ab = 2c^2$$

Therefore,

$$b^3 = 2c^3$$

From this,

$$b = \sqrt[3]{2}c = \alpha c$$

Now we conclude that

$$\alpha = \frac{b}{c} \in \mathbb{Q},$$

which is a contradiction.

We considered all possible cases when  $c \neq 0$ . Therefore, we must have that  $c = 0$ .

If  $c = 0$ , then the linear relation becomes

$$b\alpha + a = 0$$

If  $b \neq 0$ , then

$$\alpha = -\frac{a}{b} \in \mathbb{Q},$$

which is a contradiction. Therefore, we must have that  $b = 0$ .

If  $b = 0$ , then  $a = 0$ , so we only have a trivial solution  $a = b = c = 0$ .

Finally, we see that only the case  $a = b = c = 0$  is possible, so  $(1, \alpha, \alpha^2)$  is linearly independent over  $\mathbb{Q}$ .



(b)

Denote

$$F = \{a + b\alpha + c\alpha^2 \mid a, b, c \in \mathbb{Q}\}$$

Addition makes  $F$  into the abelian group.

1. If  $x, y \in F$ , then  $x = a_x + b_x\alpha + c_x\alpha^2$  and  $y = a_y + b_y\alpha + c_y\alpha^2$ , for some rational numbers  $a_x, b_x, c_x, a_y, b_y, c_y$ . Thus,

$$x + y = (a_x + a_y) + (b_x + b_y)\alpha + (c_x + c_y)\alpha^2 \in F$$

2. Associativity trivially follows, because addition is associative in  $\mathbb{C}$ , and  $F \subseteq \mathbb{C}$  (with the same law for addition).
3. Notice that  $0 = 0 + 0 \cdot \alpha + 0 \cdot \alpha^2 \in F$ . Clearly

$$x + 0 = 0 + x = x$$

for all  $x \in F$ . Thus, 0 is the additive identity in  $F$ .

4. Commutativity trivially follows, because addition is commutative in  $\mathbb{C}$ , and  $F \subseteq \mathbb{C}$  (with the same law for addition).
5. For any  $x = a + b\alpha + c\alpha^2$ , notice that  $-a + (-b)\alpha + (-c)\alpha^2 \in F$ , and that

$$x + (-a + (-b)\alpha + (-c)\alpha^2) = (-a + (-b)\alpha + (-c)\alpha^2) + x = 0$$

Therefore,

$$-x = -a + (-b)\alpha + (-c)\alpha^2 \in F$$

Multiplication makes the set of nonzero elements of  $F$  into an abelian group.

1. If  $x, y \in F^\times$ , then  $x = a_x + b_x\alpha + c_x\alpha^2$  and  $y = a_y + b_y\alpha + c_y\alpha^2$ , for some rational numbers  $a_x, b_x, c_x, a_y, b_y, c_y$ . Thus,

$$xy = (a_x a_y + 2b_x c_y + 2c_x b_y) + (a_x b_y + b_x a_y + 2c_x c_y)\alpha + (a_x c_y + b_x b_y + c_x a_y) \in F$$

2. Associativity trivially follows, because multiplication is associative in  $\mathbb{C}^\times$ , and  $F^\times \subseteq \mathbb{C}^\times$  (with the same law for multiplication).
3. Notice that  $1 = 1 + 0 \cdot \alpha + 0 \cdot \alpha^2 \in F^\times$ . Clearly

$$x \cdot 1 = 1 \cdot x = x$$

for all  $x \in F^\times$ . Thus, 1 is the multiplicative identity in  $F^\times$ .

4. Commutativity trivially follows, because multiplication is commutative in  $\mathbb{C}^\times$ , and  $F^\times \subseteq \mathbb{C}^\times$  (with the same law for multiplication).
5. We must prove that each number  $x \in F^\times$  has a multiplicative inverse.

Let  $x \in F^\times$ ;  $x = a + b\alpha + c\alpha^2$ . We want to find rational numbers  $q_1, q_2, q_3$  such that

$$(a + b\alpha + c\alpha^2)(q_1 + q_2\alpha + q_3\alpha^2) = 1$$

This is equivalent to

$$(aq_1 + 2bq_3 + 2cq_2) + (aq_2 + bq_1 + 2cq_3)\alpha + (aq_3 + bq_2 + cq_1) = 1$$

This is furthermore equivalent to the system of equations

$$\begin{aligned} aq_1 + 2cq_2 + 2bq_3 &= 1 \\ bq_1 + aq_2 + 2cq_3 &= 0 \\ cq_1 + bq_2 + aq_3 &= 0 \end{aligned}$$

We can write this in the matrix form:

$$\begin{bmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Now we want to show that the above equation has only one solution. To prove that, it is sufficient to prove that  $A = \begin{bmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{bmatrix}$  is invertible, and for that, it is sufficient to prove that  $\det A \neq 0$ . To begin

$$\det A = a^3 - 6abc + 2b^3 + 4c^3$$

Suppose that  $\det A = 0$ ; that is,  $a^3 - 6abc + 2b^3 + 4c^3$ . Since  $a, b, c$  are rational numbers, we can multiply this equation by some integer to obtain the equation

$$a_1^3 - 6a_1b_1c_1 + 2b_1^3 + 4c_1^3 = 0, \quad (2)$$

where  $a_1, b_1, c_1$  are integers. Furthermore, we can assume that they are relatively prime (we simply divide the equation by their greatest common divisor if that is not the case).

Notice that  $a_1^3 = 6a_1b_1c_1 - 1 - 2b_1^3 - 4c_1^3$ , so 2 divides  $a_1^3$ . Since 2 is prime, it must also divide  $a_1$ . Thus,  $a_1 = 2k$ , for some integer  $k$ . Plugging this into (2),

$$8k^3 - 12kb_1c_1 + 2b_1^3 + 4c_1^3 = 0$$

Now divide this equation by 2:

$$4k^3 - 6kb_1c_1 + b_1^3 + 2c_1^3 = 0 \quad (3)$$

The same as before, we conclude that 2 divides  $b_1$ , so  $b_1 = 2l$ , for some integer  $l$ . Therefore, plugging this into (3), we obtain

$$4k^3 - 12klc_1 + 8b_1^3 + 2c_1^3 = 0$$

Dividing this by 2:

$$2k^3 - 6klc_1 + 4b_1^3 + c_1^3 = 0$$

Now we get that 2 divides  $c_1$ . So, 2 is a common divisor of  $a_1$ ,  $b_1$ , and  $c_1$ , which is impossible since we assumed that their greatest common divisor is 2.

Thus, our assumption that  $\det A = 0$  was wrong, so  $\det A \neq 0$ . This means that

$$A \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

has a (unique!) solution. This also means that

$$(a + b\alpha + c\alpha^2)(q_1 + q_2\alpha + q_3\alpha^2) = 1 \stackrel{\text{comm.}}{=} (q_1 + q_2\alpha + q_3\alpha^2)(a + b\alpha + c\alpha^2)$$

Thus,  $x^{-1}$  exists for every  $x \in F^\times$ .

[Distributive law](#). This trivially follows because it holds in  $\mathbb{C}$ , and  $F \subseteq \mathbb{C}$  (with the same laws for addition and multiplication).

[Conclusion](#). We conclude that  $F$  is truly a field.

## Result

(a) Use the hint provided in the exercise. The fact that  $\alpha \notin \mathbb{Q}$  can be of great help.

(b) Check properties from Definition 3.2.2.

6. a

!!!

# 4

## Chapter 4

### Section 1

1. a

Let  $A$  be a  $l \times m$  matrix and  $B$  be a matrix of order  $n \times p$  over the field  $F$ .

We have to show the assignment  $M \mapsto AMB$  is a linear transformation from  $F^{m \times n}$  to  $F^{l \times p}$ .

Let  $\lambda \in F$  and  $M_1, M_2 \in F^{m \times n}$ . We have to show

$$A(\lambda M_1 + M_2)B = \lambda(AM_1B) + (AM_2B).$$

Now,  $A(\lambda M_1 + M_2)B$

$$= \left( (A(\lambda M_1)) + (AM_2) \right) B \text{ [Matrix multiplication is right distributive over addition]}$$

$$= (A(\lambda M_1)B) + (AM_2B) \text{ [Matrix multiplication is left distributive over addition]}$$

$$= \lambda(AM_1B) + (AM_2B) \text{ [Scalar multiplication is compatible with matrix multiplication].}$$

#### Result

3 of 3

We showed that  $A(\lambda M_1 + M_2)B = \lambda(AM_1B) + (AM_2B)$  for each  $\lambda \in F$  and each  $M_1, M_2 \in F^{m \times n}$ .

2. a

Let  $V$  be a vector space over the field  $F$ . Let  $v_1, v_2, \dots, v_n$  be  $n$ -elements of  $V$ .

Define a map  $\phi : F^n \rightarrow V$  by  $\phi(x_1, x_2, \dots, x_n) = x_1v_1 + x_2v_2 + \dots + x_nv_n$  for all  $(x_1, x_2, \dots, x_n) \in F^n$ .

We have to show  $\phi$  is a linear transformation.

So let  $(x_1, x_2, \dots, x_n)$  and  $(y_1, y_2, \dots, y_n)$  are two elements of  $F^n$  and  $c \in F$ . Then

$$\begin{aligned}\phi\left(c(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)\right) &= \phi\left((cx_1 + y_1, cx_2 + y_2, \dots, cx_n + y_n)\right) \\ &= (cx_1 + y_1)v_1 + (cx_2 + y_2)v_2 + \dots + (cx_n + y_n)v_n \\ &= ((cx_1)v_1 + (cx_2)v_2 + \dots + (cx_n)v_n) + (y_1v_1 + y_2v_2 + \dots + y_nv_n) \\ &= c(x_1v_1 + x_2v_2 + \dots + x_nv_n) + (y_1v_1 + y_2v_2 + \dots + y_nv_n) \\ &= c\phi(x_1, x_2, \dots, x_n) + \phi(y_1, y_2, \dots, y_n).\end{aligned}$$

Therefore  $\phi$  is a linear transformation.

## Result

3 of 3

We showed that  $\phi\left(c(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)\right) = c\phi(x_1, x_2, \dots, x_n) + \phi(y_1, y_2, \dots, y_n)$  for all  $c \in F$  and for all  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in F^n$ .

## 3. a

Let  $A$  be a  $m \times n$  matrix over a field  $F$ .

We have to use dimension formula to show that the space of solutions of system of linear equations  $AX = 0$

has dimension at least  $n - m$ .

Let us consider the function  $\phi : F^n \rightarrow F^m$  defined by  $\phi(X) = AX, \forall X \in F^n$ . The map  $\phi$  is linear : for  $c \in F$  and  $X_1, X_2 \in F^n$  we have  $\phi(cX_1 + X_2) = A(cX_1 + X_2) = A(cX_1) + AX_2 = cAX_1 + AX_2$ , using the fact that matrix multiplication is right distributive over addition and scalar multiplication is compatible with matrix multiplication.

Now  $\ker(\phi) = \{X \in F^n : \phi(X) = 0\}$  = solution of system of linear equations  $AX = 0$ . Now dimension formula says,  $\dim(\ker(\phi)) + \dim(\text{im}(\phi)) = \dim(F^n) = n$ . Since  $\text{im}(\phi)$  is a subspace of  $F^m$ , so dimension of  $\text{im}(\phi)$  can be at most dimension of  $F^m$ . Therefore,  $\dim(\ker(\phi)) = n - \dim(\text{im}(\phi)) \geq n - \dim(F^m) = n - m$ .

## Result

3 of 3

We used the dimension formula to the linear transformation  $\phi : F^n \rightarrow F^m$  defined by  $\phi(X) = AX, \forall X \in F^n$ .

## 4. a

Let  $A$  be a  $m \times n$  matrix over the field  $F$  of rank 1. We have to show  $A$  can be written as  $A = XY^t$ ,

where  $X$  and  $Y$  are  $m$  and  $n$  dimensional column vectors.

Write  $A$  as  $A = [A_1 : A_2 : \cdots : A_n]$ , where  $A_1, A_2, \dots, A_n$  are columns of  $A$ . Since  $\text{rank}(A) = 1$  we have  $j \in \{1, \dots, n\}$  such that  $A_j \neq 0$  and some scalars  $c_1, \dots, c_{j-1}, c_{j+1}, \dots, c_n \in F$  with

$$A = [c_1 A_j : c_2 A_j : \cdots : c_{j-1} A_j : A_j : c_{j+1} A_j : \cdots : c_n A_j].$$

$$= A_j \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{j-1} \\ 1 \\ c_{j+1} \\ \vdots \\ c_n \end{bmatrix}^t.$$

So our  $X = A_j$  and

$$Y = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{j-1} \\ 1 \\ c_{j+1} \\ \vdots \\ c_n \end{bmatrix}^t$$

Note that we have to choose a non-zero column of  $A$

and correspondingly we have unique scalar multiplies of this column to write other columns of  $A$ .

## Result

3 of 3

Writing  $A$  as  $A = [c_1 A_j : c_2 A_j : \cdots : c_{j-1} A_j : A_j : c_{j+1} A_j : \cdots : c_n A_j]$  for some non-zero column  $A_j$  of  $A$  and for some scalars  $c_1, \dots, c_{j-1}, c_{j+1}, \dots, c_n \in F$  we have

$$A = A_j \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{j-1} \\ 1 \\ c_{j+1} \\ \vdots \\ c_n \end{bmatrix}^t.$$

5. a



Let  $U, W$  be two vector spaces over a field  $F$ . We have to show the operations

$$(u, w) + (u', w') = (u + u', w + w'); \forall u, u' \in U \text{ \& } w, w' \in W,$$

$$c(u, w) = (cu, cw), \forall c \in F \text{ \& } \forall u \in U \text{ \& } \forall w \in W$$

make the product space  $U \times W$  into a vector space.

Let  $0_U$  and  $0_W$  be the identities of  $U$  and  $W$  respectively. Then for  $u, u' \in U$  and  $w, w' \in W$  we have  $u + u' = u' + u$  and  $w + w' = w' + w$  and  $0_U + u = u = u + 0_U$  and  $0_W + w = w = w + 0_W$  so that,  $(u, w) + (u', w') = (u + u', w + w') = (u' + u, w' + w) = (u', w') + (u, w)$  and  $(u, w) + (0_U, 0_W) = (u + 0_U, w + 0_W) = (u, w) = (0_U + u, 0_W + w) = (0_U, 0_W) + (u, w)$ . So the  $+$  operation makes  $U \times W$  into a commutative group with identity  $(0_U, 0_W)$ .

Now  $1(u, w) = (1u, 1w) = (u, w), \forall u \in U, \forall w \in W$ .

Next let  $a, b \in F$ , then for  $u \in U$  and  $w \in W$  we have  $(ab)(u, w) = ((ab)u, (ab)w) = (a(bu), a(bw)) = a(bu, bw) = a(b(u, w))$ . This proves the associative law of scalar multiplication. Similarly, we have  $(a + b)(u, w) = ((a + b)u, (a + b)w) = (au + bu, aw + bw) = (au, aw) + (bu, bw) = a(u, w) + b(u, w)$ . This shows distributive law of scalar multiplication.

Let  $U$  and  $W$  be subspaces of a vector space  $V$ . Define a map  $T : U \times W \rightarrow V$  by  $T(u, w) = u + w, \forall u \in U, \forall w \in W$ . We have to show  $T$  is linear.

So let  $(u_1, w_1), (u_2, w_2)$  are two elements of  $U \times W$  and  $c$  be a scalar. So that,

$$\begin{aligned} T(c(u_1, w_1) + (u_2, w_2)) &= T((cu_1, cw_1) + (u_2, w_2)) \\ &= T(cu_1 + w_1, cu_2 + w_2) = (cu_1 + w_1) + (cu_2 + w_2) \\ &= c(u_1 + w_1) + (u_2 + w_2) = cT(u_1, w_1) + T(u_2, w_2). \end{aligned}$$

So  $T$  is linear.

Let  $U$  and  $W$  be subspaces of a finite dimensional vector space  $V$ . Define a map  $T : U \times W \rightarrow V$  by  $T(u, w) = u + w, \forall u \in U, \forall w \in W$ . Now dimension formula says,  $\dim(\text{range}(T)) + \dim(\text{kernel}(T)) = \dim(U \times W)$ .

Note that,  $\text{range}(T) = U + W$ . So  $\dim(\text{range}(T)) = \dim(U) + \dim(W) - \dim(U \cap W)$ .

Also,  $\ker(T) = \{(u, w) \in U \times W : u = -w\}$ . Now consider a basis  $\mathcal{B}$  of  $U \cap W$ . Consider the set  $\mathcal{C} := \{(u, -u) : u \in \mathcal{B}\}$ . Certainly,  $\mathcal{C}$  is a linearly independent subset of the vector space  $U \times W$ :

$\sum_{k=1}^n \lambda_k(u_k, -u_k) = (0, 0)$  for scalars  $\lambda_1, \dots, \lambda_n$  implies  $(\sum_{k=1}^n \lambda_k u_k, -\sum_{k=1}^n \lambda_k u_k) = (0, 0)$  i.e.

$\sum_{k=1}^n \lambda_k u_k = 0$  i.e.  $\lambda_1 = 0, \dots, \lambda_n = 0$ . Next for any  $(u, -u) \in \ker(T)$  we have  $u \in U \cap W$ , so we have

scalars  $c_1, \dots, c_m$  and vectors  $u'_1, \dots, u'_m \in \mathcal{B}$  such that  $u = \sum_{k=1}^m c_k u'_k$ . Hence,  $(u, -u) =$

$(\sum_{k=1}^m c_k u'_k, -\sum_{k=1}^m c_k u'_k) = \sum_{k=1}^m c_k(u'_k, -u'_k)$ . Hence,  $\mathcal{C}$  is a basis of  $\ker(T)$ . But  $\dim(\ker(T)) =$  cardinality of  $\mathcal{C} =$  cardinality of  $\mathcal{B} = \dim(U \cap W)$ .

Next let,  $\{x_1, \dots, x_r\}$  be a basis of  $U$  and  $\{y_1, \dots, y_s\}$  is a basis of  $W$ . Let  $\{\mu_1, \dots, \mu_{r+s}\}$  be a set of scalars with  $\sum_{k=1}^r \mu_k(x_k, 0) + \sum_{k=1}^s \mu_{k+r}(0, y_k) = (0, 0)$  i.e.  $\sum_{k=1}^r \mu_k x_k = 0 = \sum_{k=1}^s \mu_{k+r} y_k$ . Hence,  $0 = \mu_1 = \dots = \mu_{r+s}$ . Now choose  $(u, w) \in U \times W$  then we have scalars  $\alpha_1, \dots, \alpha_r$  and scalars  $\beta_1, \dots, \beta_s$  such that  $u = \sum_{k=1}^r \alpha_k x_k$  and  $w = \sum_{k=1}^s \beta_k y_k$ . Then  $\sum_{k=1}^r \alpha_k(x_k, 0) + \sum_{k=1}^s \beta_k(0, y_k) = (u, w)$ . Therefore,  $\{(x_1, 0), \dots, (x_r, 0)\} \cup \{(0, y_1), \dots, (0, y_s)\}$  is a basis for  $U \times W$ . That's  $\dim(U \times W) = r + s = \dim(U) + \dim(W)$ .

Using the above discussion and the dimension formula  $\dim(\text{range}(T)) + \dim(\text{kernel}(T)) = \dim(U \times W)$  we have the identity,

$$\{\dim(U) + \dim(W) - \dim(U \cap W)\} + \dim(U \cap W) = \dim(U) + \dim(W).$$



We showed that  $\dim(\text{im}(T)) = \dim(U) + \dim(W) - \dim(U \cap W)$  and  $\dim(\ker(T)) = \dim(U \cap W)$   
and  $\dim(U \times W) = \dim(U) + \dim(W)$ .

## Section 2

1. a

Let  $A$  and  $B$  are two  $2 \times 2$  matrices over  $F$ . We have to find the matrix of the operator  $T : F^{2 \times 2} \rightarrow F^{2 \times 2}$  defined by  $T(M) = AMB$ ,

$\forall M \in F^{2 \times 2}$  w.r.t. to the basis  $e_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $e_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ ,  $e_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ ,  $e_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$  of  $F^{2 \times 2}$ .

Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

. Then

$$\begin{aligned} T(e_{11}) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{bmatrix} = (a_{11}b_{11})e_{11} + (a_{11}b_{12})e_{12} + (a_{21}b_{11})e_{21} + (a_{21}b_{12})e_{22}. \end{aligned}$$

Next

$$\begin{aligned} T(e_{12}) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & a_{11} \\ 0 & a_{21} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{21} & a_{11}b_{22} \\ a_{21}b_{21} & a_{21}b_{22} \end{bmatrix} = (a_{11}b_{21})e_{11} + (a_{11}b_{22})e_{12} + (a_{21}b_{21})e_{21} + (a_{21}b_{22})e_{22}. \end{aligned}$$

And

$$\begin{aligned} T(e_{21}) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{12} & 0 \\ a_{22} & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{12}b_{11} & a_{12}b_{12} \\ a_{22}b_{11} & a_{22}b_{12} \end{bmatrix} = (a_{12}b_{11})e_{11} + (a_{12}b_{12})e_{12} + (a_{22}b_{11})e_{21} + (a_{22}b_{12})e_{22}. \end{aligned}$$

And

$$\begin{aligned} T(e_{22}) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & a_{12} \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{bmatrix} = (a_{12}b_{21})e_{11} + (a_{12}b_{22})e_{12} + (a_{22}b_{21})e_{21} + (a_{22}b_{22})e_{22}. \end{aligned}$$

Therefore, the matrix of  $T$  w.r.t. the basis  $\{e_{11}, e_{12}, e_{21}, e_{22}\}$  is

$$\begin{bmatrix} a_{11}b_{11} & a_{11}b_{21} & a_{12}b_{11} & a_{12}b_{21} \\ a_{11}b_{12} & a_{11}b_{22} & a_{12}b_{12} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{21} & a_{22}b_{11} & a_{22}b_{21} \\ a_{21}b_{12} & a_{21}b_{22} & a_{22}b_{12} & a_{22}b_{22} \end{bmatrix}.$$

## Result

Letting

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

we showed that matrix of  $T : M \mapsto AMB$  w.r.t. the basis  $\{e_{11}, e_{12}, e_{21}, e_{22}\}$  is

$$\begin{bmatrix} a_{11}b_{11} & a_{11}b_{21} & a_{12}b_{11} & a_{12}b_{21} \\ a_{11}b_{12} & a_{11}b_{22} & a_{12}b_{12} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{21} & a_{22}b_{11} & a_{22}b_{21} \\ a_{21}b_{12} & a_{21}b_{22} & a_{22}b_{12} & a_{22}b_{22} \end{bmatrix}.$$

## 2. a

Let  $A$  be an  $n \times n$  matrix and let  $V$  denote the vector space of  $n$ -dimensional row vectors. Note that an  $n \times n$  matrix can be composed with a row vector only from the right. If we do that, then it is easy to see that

$$T(v) = vA$$

is indeed a linear operator. We will find its matrix in the standard basis. The standard basis is:

$$\begin{aligned} e_1 &= (1 \ 0 \ \dots \ 0) \\ e_2 &= (0 \ 1 \ \dots \ 0) \\ &\dots \\ e_n &= (0 \ 0 \ \dots \ 1) \end{aligned}$$

(Note that these are  $1 \times n$  matrices i.e. row vectors.) Consider how right multiplication acts on  $e_k$ :

$$\begin{aligned} e_k A &= [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \\ &= [a_{k1} \ \dots \ a_{k \ k-1} \ a_{kk} \ a_{k \ k+1} \ \dots \ a_{kn}] \\ &= a_{k1}e_1 + \dots + a_{kn}e_n \end{aligned}$$

This allows us to write the matrix of the operator  $T$  in the standard basis. The columns will be the coefficients above, so:

$$\begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} = A^t$$

Thus the "right multiplication by  $A$ " operator has  $A^t$  as its matrix in the standard basis.

## Result

2 of 2

Acting on the standard basis with right multiplication by  $A$ , we find the matrix of this operator. It turns out to be  $A^t$ .

### 3. a

We have to find all real  $2 \times 2$  matrices which carry the line  $y = x$  to  $y = 3x$ .

Note that, each point on the line  $y = x$  can be written as  $(t, t)$  for some  $t \in \mathbb{R}$  and each point of the line  $y = 3x$  can be written as  $(s, 3s)$  for some  $s \in \mathbb{R}$ .

Let

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be a matrix which takes the line  $y = x$  to  $y = 3x$ . Now note that,

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is on the line  $y = x$ . So that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} s \\ 3s \end{bmatrix}$$

for some  $s \in \mathbb{R}$ . That's  $a + b = s$  and  $c + d = 3s$  i.e.  $3(a + b) = c + d$ .

**Check:—**For the point

$$\begin{bmatrix} t \\ t \end{bmatrix}$$

on the line  $y = x$  we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} t \\ t \end{bmatrix} = t \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = t \begin{bmatrix} s \\ 3s \end{bmatrix} = \begin{bmatrix} ts \\ 3ts \end{bmatrix},$$

which is certainly on the line  $y = 3x$ .

## Result

The required matrix has the following form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

with  $3(a + b) = c + d$ .

#### 4. a

Let  $A$  be an  $m \times n$  matrix. The rank is defined as the dimension of the image of the corresponding operator. Let  $e_1, \dots, e_n$  be the standard basis (for  $F^n$ ). Since these vectors span the domain of  $A$ ,

$$Ae_1, Ae_2, \dots, Ae_n$$

must span all of the image of  $A$ . The number  $r$  of linearly independent vectors among these is the rank. Note that these vectors are just the columns of the matrix  $A$ . Sort the above vectors so that the first  $r$  are linearly independent (if the matrix is non-zero there must be some such vector). Write them as

$$v_1, v_2, \dots, v_r, v_{r+1}, \dots, v_n$$

Switching rows or columns is a matrix operation which can be realized by multiplication with the appropriate matrix. Switching columns is done by multiplication from the right. Suppose this was achieved by a matrix  $P_1$ . We have arrived at the matrix

$$AP_1 = [v_1 \ v_2 \ \dots \ v_n]$$

whose first  $r$  columns make a linearly independent set.

Now, since there are only  $r$  linearly independent vectors, we must have a relation

$$v_{r+1} = c_1 v_1 + c_2 v_2 + \dots + c_r v_r$$

But then adding the first column  $-c_1$  times, adding the second column  $-c_2$  and so on, to the  $r+1$ -st column, we would arrive at a matrix which has for the  $r+1$ -st column the zero vector. The same goes for  $v_{r+2}, \dots, v_n$ .

Adding columns (multiplied by a number) is an invertible matrix operation which can be realized by multiplying with the appropriate matrices from the right. Thus there is a  $P_2$  such that

$$AP_1 P_2 = [v_1 \ v_2 \ \dots \ v_r \ 0 \ \dots \ 0]$$

Now, consider the matrix of the form

$$A' = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where  $I_r$  is shorthand for the  $r \times r$  identity matrix and the rest of the entries are all zero. Let  $c_1, \dots, c_r$  be arbitrary. By multiplying the first row  $c_1$  times, adding the second row  $c_2$  times to the first row and so on, we can arrive at the matrix with first row:

$$\begin{bmatrix} c_1 & c_2 & \dots & c_r & 0 & \dots & 0 \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{bmatrix}$$

Thus we can make the first row into an arbitrary vector with the first  $r$  elements non-zero. The same can be accomplished for any row. Thus we conclude that by row operations, we can get an arbitrary matrix of the form:

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1r} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & b_{2r} & 0 & \dots & 0 \\ & & & & & & \\ & & & & & & \\ b_{m1} & b_{m2} & \dots & b_{mr} & 0 & \dots & 0 \end{bmatrix}$$

Since only row operations were used, we conclude that there is an invertible matrix  $Q$  such that  $QA'$  is equal to the above matrix.

Finally, write  $P = P_1 P_2$ . This is an invertible matrix since  $P_1, P_2$  were. Note that  $AP$  has exactly the form as the matrix  $QI_r$  in the previous paragraph. Thus for a suitable choice of  $Q$ , we would have

$$AP = QA'$$

Multiplying with the inverse of  $Q$  from the left, we finally arrive at the fact that there are invertible matrices  $P, Q$  such that

$$A' = Q^{-1}AP$$

## Result

5 of 5

The rank  $r$  is the number of linearly independent column vectors. Reorder the columns of the matrix  $A$  so that the first  $r$  vectors form a linearly independent set.

The other vectors can be expressed with them, thus for an appropriate invertible matrix  $P$ , column operations give a matrix

$$AP = [v_1 \ v_2 \ \dots \ v_r \ 0 \ \dots \ 0]$$

Finally, consider  $A'$ . With appropriate row operations, this can be converted to any matrix whose columns are 0 after the  $r$ -th one. Thus for an appropriate  $Q$ , we have  $QA' = AP$ , as desired.

5. a

Let  $A = [a_{ij}]_{i=1, j=1}^{i=m, j=n}$  be a  $m \times n$  matrix of rank  $r$ .

Let  $I = \{i_1, i_2, \dots, i_r\} \subseteq \{1, 2, \dots, m\}$  and  $J = \{j_1, j_2, \dots, j_r\} \subseteq \{1, 2, \dots, n\}$  such that  $i_1 < i_2 < \dots < i_r$  and  $j_1 < j_2 < \dots < j_r$ .

Also let,  $\{A_{i_1*}, A_{i_2*}, \dots, A_{i_r*}\}$  be a linearly independent subset of rows of  $A$  and  $\{A_{*j_1}, A_{*j_2}, \dots, A_{*j_r}\}$  be a linearly independent subset of columns of  $A$ .

We have to show the matrix  $M := [a_{i_k j_l}]_{k=1, l=1}^{k=r, l=r}$  is invertible.

Consider the matrix  $B = [A_{*j_1} : A_{*j_2} : \dots : A_{*j_r}]$ . Now column rank of  $B$  is  $r$ . Hence row rank of  $B$  is  $r$  also. Let  $B_{l_1*}, B_{l_2*}, \dots, B_{l_r*}$  be a independent set of rows of  $B$ , where  $\{l_1, \dots, l_r\} \subseteq \{1, \dots, m\}$  and  $l_1 < l_2 < \dots < l_r$ . Hence  $A_{l_1*}, A_{l_2*}, \dots, A_{l_r*}$  is also a independent set of rows of  $A$ . Since  $\{A_{i_1*}, A_{i_2*}, \dots, A_{i_r*}\}$  is a basis of row space of  $A$ , each  $A_{l_t*}$  can be written as linear combination of  $A_{i_1*}, A_{i_2*}, \dots, A_{i_r*}$  for  $t = 1, 2, \dots, r$ . Hence, each  $B_{l_t*}$  can be written as linear combination of rows of  $M$  for  $t = 1, 2, \dots, r$ . That is row space of  $B$  is same as row space of  $M$ . Hence row rank of  $B$  is same as row rank of  $M$ . But row rank of  $B$  is  $r$  and  $M$  is a  $r \times r$  matrix. Hence  $M$  is invertible.

## Result

3 of 3

We use the fact that for a  $m \times n$  matrix row rank is same as column rank.

# Section 3



1. a

Consider the linear operator  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  defined by  $T(x_1, \dots, x_n)^t = (x_1 + x_n, x_2 + x_{n-1}, \dots, x_n + x_1)^t$ .

We have to find kernel and image of this operator.

Suppose,  $(x_1, x_2, \dots, x_n)^t \in \ker(T)$ , then  $T(x_1, x_2, \dots, x_n)^t = (0, 0, \dots, 0)^t$  i.e.  $(x_1 + x_n, x_2 + x_{n-1}, \dots, x_n + x_1)^t = (0, 0, \dots, 0)^t$ . That is,  $x_n = -x_1, x_{n-1} = -x_2, \dots$ . Conversely if  $(a_1, a_2, \dots, a_n)^t \in \mathbb{R}^n$  is such that  $a_k = -a_{n-k}, \forall k = 1, 2, \dots, (n-1)$ , then  $(a_1, \dots, a_n)^t \in \ker(T)$ .

Therefore,  $\ker(T) = \{(x_1, \dots, x_n)^t \in \mathbb{R}^n : x_k = -x_{n-k}, \forall k = 1, 2, \dots, (n-1)\}$ .

Now suppose  $(y_1, \dots, y_n)^t \in \text{im}(T)$ , then there is  $(x_1, \dots, x_n) \in \mathbb{R}^n$  such that  $(y_1, y_2, \dots, y_n)^t = (x_1 + x_n, x_2 + x_{n-1}, \dots, x_n + x_1)^t$ . That is  $y_k = x_{n-k} + x_k, \forall k = 1, 2, \dots, (n-1)$ . Conversely, let  $(b_1, b_2, \dots, b_n)^t \in \mathbb{R}^n$  be such that  $b_k = b_{n-k}, \forall k = 1, 2, \dots, (n-1)$ . Then  $T(\frac{b_1}{2}, \frac{b_2}{2}, \frac{b_3}{2}, \dots, \frac{b_{n-1}}{2}, \frac{b_n}{2})^t = (b_1, b_2, b_3, \dots, b_{n-1}, b_n)^t$ .

Hence  $\text{im}(T) = \{(y_1, y_2, \dots, y_n)^t \in \mathbb{R}^n : y_k = y_{n-k}, \forall k = 1, 2, \dots, (n-1)\}$ .

## Result

3 of 3

We show  $\ker(T) = \{(x_1, \dots, x_n)^t \in \mathbb{R}^n : x_k = -x_{n-k}, \forall k = 1, 2, \dots, (n-1)\}$  and  $\text{im}(T) = \{(y_1, y_2, \dots, y_n)^t \in \mathbb{R}^n : y_k = y_{n-k}, \forall k = 1, 2, \dots, (n-1)\}$ .

2. a

- [a] Consider the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

where  $c$  is non-zero. We will show that we can change the  $a$  into 0 by conjugation.

Since  $c$  is the only element we are sure is non-zero, we should use it to eliminate  $a$ . We could do this by a row operation. Indeed, for

$$S = \begin{bmatrix} 1 & -\frac{a}{c} \\ 0 & 1 \end{bmatrix}$$

one will find that

$$SA = \begin{bmatrix} 1 & -\frac{a}{c} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & b - \frac{ad}{c} \\ c & d \end{bmatrix}$$

Now, note that multiplication by  $S^{-1}$  from the right will be a column operation, but this will not change  $SA$  from having a 0 term. Indeed,

$$S^{-1} = \begin{bmatrix} 1 & \frac{a}{c} \\ 0 & 1 \end{bmatrix}$$

and

$$(SA)S^{-1} = \begin{bmatrix} 0 & b - \frac{ad}{c} \\ c & d \end{bmatrix} \begin{bmatrix} 1 & \frac{a}{c} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & * \\ * & * \end{bmatrix}$$

Thus for the above choice of the elementary matrix  $S$ , we find that  $SA S^{-1}$  has 0 instead of the element  $a$ .

- [b)] Suppose  $c = 0$ . If  $b \neq 0$ , then we can proceed similarly as above by choosing the appropriate  $S$ :

$$S^{-1}AS = \begin{bmatrix} 1 & 0 \\ \frac{a}{b} & 1 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{a}{b} & 1 \end{bmatrix} = \begin{bmatrix} 0 & * \\ * & * \end{bmatrix}$$

We see that  $S^{-1}AS$  has 0 instead of  $a$  in the upper left corner. We conclude that even if  $b \neq 0$ , the matrix is similar to a matrix with 0 instead of  $a$ .

Now, assume that  $b = c = 0$ . Then for a matrix

$$R = \begin{bmatrix} 1 & -x \\ 0 & 1 \end{bmatrix}$$

, we have:

$$R^{-1}AR = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \begin{bmatrix} 1 & -x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & x(a-d) \\ 0 & a \end{bmatrix}$$

for any  $x$ . If  $a - d \neq 0$ , then  $R^{-1}AR$  can be transformed to a matrix with 0 instead of  $a$  just like above. Since similarity is transitive, we conclude that  $A$  is again similar to a matrix with 0 instead of  $a$ .

Now, if we were to have the case that  $b = c = 0$  and  $a = d$ , then we would have that  $A = aI$ , where  $I$  is the identity matrix. Obviously then  $S^{-1}AS = A$ , so that it cannot be changed to anything else.

We conclude that the matrix  $A$  is similar to a matrix with 0 instead of  $a$  **if and only if**  $b \neq 0$  or  $a \neq d$  (or if  $A$  is the zero matrix).

## Result

3 of 3

- a) By choosing the appropriate elementary similarity matrix

$$S = \begin{bmatrix} 1 & -\frac{a}{c} \\ 0 & 1 \end{bmatrix}$$

, one easily shows this explicitly.

d

- b) Similar to a), if  $b \neq 0$ , one can turn the 'a' term into 0. If  $b = 0$ , then one can still do that if  $a \neq d$ . But if even this is not met, then it is possible only if  $A$  is the zero matrix.

## 3. a

Let  $T$  be a linear operator on the vector space  $V$  with  $\dim(V) = 2$  such that  $T$  is not multiplication by a scalar. We have to show there is  $v \in V$  such that  $\{v, T(v)\}$  is a basis of  $V$ .

In particular,  $T$  is not zero operator. Hence there is a  $w \in V - \{0\}$  such that  $T(w) \neq 0$ . Now if  $\{w, T(w)\}$  is linearly independent then we are done.

Next consider the case when  $\{w, T(w)\}$  is linearly dependent i.e. we have two scalars  $\lambda$  and  $\mu$ , not both zero such that  $\lambda w + \mu T(w) = 0$ . Actually  $\lambda \neq 0$  and  $\mu \neq 0$  as  $T(w) \neq 0$  and  $w \neq 0$ . Choose  $u \in V$  such that  $\{u, w\}$  is a basis of  $V$ . Write  $T(u) = \delta w + \alpha u$  for some scalars  $\delta$  and  $\alpha$ . Let  $v = \gamma w + u$  where  $\gamma$  is a scalar such that  $\gamma = 0$  if  $\delta \neq 0$  and  $\gamma \neq 0$  if  $\delta = 0$ . Note that  $\delta = 0$  implies  $\alpha + \lambda\mu^{-1} \neq 0$  as  $T$  is not multiplication by a fixed scalar. Now let  $c_1$  and  $c_2$  are two scalars, not both zero such that  $c_1 v + c_2 T(v) = 0$  i.e.  $c_1(\gamma w + u) + c_2(\gamma T(w) + \delta w + \alpha u) = 0$  i.e.  $c_1(\gamma w + u) + c_2(-\gamma\lambda\mu^{-1}w + \delta w + \alpha u) = 0$  i.e.  $(c_1 + c_2\alpha)u + (c_1\gamma - c_2\gamma\lambda\mu^{-1} + c_2\delta)w = 0$ . Now,  $\{u, w\}$  is linearly independent implies  $c_1 + c_2\alpha = 0 = c_1\gamma - c_2\gamma\lambda\mu^{-1} + c_2\delta$ . Therefore,  $c_2(-\alpha\gamma - \gamma\lambda\mu^{-1} + \delta) = 0$ . Now  $c_2 \neq 0$  as  $v \neq 0$ . That is  $\gamma(\alpha + \lambda\mu^{-1}) = -\delta$ . By our choice of  $\gamma$  this leads to a contradiction. Therefore,  $\{v, T(v)\}$  is a linearly independent set. Hence we are done.



Now we have to find the matrix of  $T$  w.r.t the basis  $\{v, T(v)\}$ .

Note that,  $T(T(v)) = av + bT(v)$  for some scalars  $a, b$ . Therefore the matrix of  $T$  w.r.t. the basis  $\{v, T(v)\}$  is  $\begin{bmatrix} 0 & a \\ 1 & b \end{bmatrix}$ .

## Result

3 of 3

Matrix of  $T$  w.r.t. the basis  $\{v, T(v)\}$  is

$$\begin{bmatrix} 0 & a \\ 1 & b \end{bmatrix}$$

for some scalars  $a$  and  $b$ .

4. a

We show that, for a  $n \times n$  complex matrix  $B$  the operator  $T$  on the space of all  $n \times n$  matrices defined by  $T(A) = AB - BA$  is singular.

## Step 2

2 of 3

Certainly,  $B = 0$  implies  $T = 0$ . So we are done in this case as  $\ker(T) \neq \{0\}$ .

Next let  $B \neq 0$ , then  $B \in \ker(T)$  as  $(B)B - B(B) = B^2 - B^2 = 0$  i.e.  $\ker(T) \neq \{0\}$  in this case also.

Hence in either case  $T$  is singular.

## Result

3 of 3

$B = 0$  implies  $\ker(T)$  is the whole space and  $B \neq 0$  implies  $\ker(T)$  contains the non-zero vector  $B$ . That is in either case  $T$  is singular.

# Section 4

1. a

Let  $T$  be a linear operator on a vector space  $V$  and  $\lambda$  be a scalar.

Let  $V^{(\lambda)}$  be the set of all eigenvectors of  $T$  corresponding to the eigenvalue  $\lambda$ ,

together with 0-vector. We have to show  $V^{(\lambda)}$  is an invariant subspace of  $T$ .

Note that,  $V^{(\lambda)}$  is non-empty as it always contains 0. Now if  $v_1, v_2 \in V^{(\lambda)}$  we have  $Tv_1 = \lambda v_1$  and  $Tv_2 = \lambda v_2$ . Now let  $c$  be a scalar then  $T(v_1 + cv_2) = T(v_1) + T(cv_2) = T(v_1) + cT(v_2) = \lambda v_1 + c(\lambda v_2) = \lambda(v_1 + cv_2)$ . That's  $V^{(\lambda)}$  is a vector space.

Now let  $v \in V^{(\lambda)}$  then  $Tv = \lambda v$  i.e.  $T(Tv) = T(\lambda v) = \lambda T(v)$  i.e.  $T(v) \in V^{(\lambda)}$ . Since  $v \in V^{(\lambda)}$  is arbitrary we have  $T(V^{(\lambda)}) \subseteq V^{(\lambda)}$ . Hence  $V^{(\lambda)}$  is an invariant subspace of  $V$  under  $T$ .

## Result

3 of 3

We showed that,  $T(V^{(\lambda)}) \subseteq V^{(\lambda)}$ .

## 2. a

Let  $T$  be a linear operator on a vector space  $V$  over a field  $F$  with  $\text{char}(F) \neq 2$ . Also, let  $T^2 = I$ . Now, let  $v \in V$  and suppose  $v - Tv$  is a non-zero vector. Then  $T(v - Tv) = Tv - T^2v = Tv - Iv = Tv - v = (-1)(v - Tv)$ . Since,  $\text{char}(F) \neq 0$ , we can say  $2 := 1 + 1 \neq 0$ , so that 2 is an invertible element in  $F$ . Hence, we have  $T(\frac{1}{2}(v - Tv)) = (-1)(\frac{1}{2}(v - Tv))$ . Therefore,  $\frac{1}{2}(v - Tv)$  is an eigenvector of  $T$  with eigenvalue  $-1$ .

Similarly, for any  $w \in V$  with  $(w + Tw) \neq 0$  implies  $T(\frac{1}{2}(w + Tw)) = \frac{1}{2}Tw + \frac{1}{2}T^2w = \frac{1}{2}Tw + \frac{1}{2}Iw = \frac{1}{2}Tw + \frac{1}{2}w = (+1)(\frac{1}{2}(w + Tw))$ . Therefore,  $\frac{1}{2}(w + Tw)$  is an eigenvector of  $T$  with eigenvalue  $+1$ .

Now, note that  $z = \frac{1}{2}(z - Tz) + \frac{1}{2}(z + Tz)$  for any  $z \in V$ . So, when  $z \neq 0$  either  $\frac{1}{2}(z - Tz) \neq 0$  or  $\frac{1}{2}(z + Tz) \neq 0$ , i.e.  $V = V^{(+1)} + V^{(-1)}$ , where  $V^{(+1)} := \{x \in V : Tx = x\}$  and  $V^{(-1)} := \{y \in V : Ty = -y\}$ . Note that  $u \in V^{(+1)} \cap V^{(-1)}$  implies  $Tu = u$  and  $Tu = -u$ , and these further imply  $u = 0$ . So, we have  $V = V^{(+1)} \oplus V^{(-1)}$ . Therefore,  $T$  is diagonalizable.

Let  $T$  be a linear operator on complex vector space  $V$  with  $T^4 = I$ . Now

$$4 \cdot I = [I + T^3 + T^2 + T] + [I - T^3 + T^2 - T] + [I - iT^3 - T^2 + iT] + [I + iT^3 - T^2 - iT].$$

Also,

$$\begin{aligned} T \circ (I + T^3 + T^2 + T) &= T + T^4 + T^3 + T^2 = I + T^3 + T^2 + T, \\ T \circ (I - T^3 + T^2 - T) &= T - I + T^3 - T^2 = -(I - T^3 + T^2 - T), \\ T \circ (I - iT^3 - T^2 + iT) &= T - iI - T^3 + iT^2 = -i(I - iT^3 - T^2 + iT), \\ T \circ (I + iT^3 - T^2 - iT) &= T + iI - T^3 - iT^2 = i(I + iT^3 - T^2 - iT). \end{aligned}$$

Write,  $T_1 = I + T^3 + T^2 + T$  and  $T_2 = I - T^3 + T^2 - T$  and  $T_3 = I - iT^3 - T^2 + iT$  and  $T_4 = I + iT^3 - T^2 - iT$ .

Then for any  $v \in V$  we have  $v = \frac{1}{4}[T_1v + T_2v + T_3v + T_4v]$  and  $T_1v \in V^{(1)}$ ,  $T_2v \in V^{(-1)}$ ,  $T_3v \in V^{(-i)}$ ,  $T_4v \in V^{(i)}$ . Hence  $V = V^{(1)} + V^{(-1)} + V^{(-i)} + V^{(i)}$ .

Now suppose,  $0 = x + y + w + z$  where  $x \in V^{(1)}$ ,  $y \in V^{(-1)}$ ,  $w \in V^{(-i)}$  and  $z \in V^{(i)}$ . Hence,  $0 = T0 = Tx + Ty + Tw + Tz = x - y - iw + iz$  and  $0 = T^20 = T^2x + T^2y + T^2w + T^2z = x + y - w - z$  and  $0 = T^30 = T^3x + T^3y + T^3w + T^3z = x - y + iw - iz$ . Solving these 4 equations we have  $0 = x = y = w = z$ . Therefore,  $V = V^{(1)} \oplus V^{(-1)} \oplus V^{(-i)} \oplus V^{(i)}$ .

## Result

3 of 3

We use the fact that,  $4I = 4 \cdot I = [I + T^3 + T^2 + T] + [I - T^3 + T^2 - T] + [I - iT^3 - T^2 + iT] + [I + iT^3 - T^2 - iT]$ , where  $T^4 = I$ .

## 3. a

Let  $T$  be a linear operator on a vector space  $V$  and suppose  $W_1$  and  $W_2$  are two  $T$ -invariant subspace of  $V$ .

We have to show  $W_1 + W_2$  and  $W_1 \cap W_2$  are also  $T$ -invariant.

Note that  $W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1 \text{ \& } w_2 \in W_2\}$ . So let  $x + y \in W_1 + W_2$  be arbitrary vector of  $W_1 + W_2$  with  $x \in W_1$  and  $y \in W_2$ . Now  $W_1$  is  $T$ -invariant implies  $T(x) \in W_1$ . Also  $W_2$  is  $T$ -invariant implies  $T(y) \in W_2$ . Therefore,  $T(x) + T(y) \in W_1 + W_2$ . Since  $x + y$  is arbitrary vector of  $W_1 + W_2$  we have  $T(W_1 + W_2) \subseteq W_1 + W_2$ .

Now let  $v \in W_1 \cap W_2$  be arbitrary. Then  $v \in W_1$  and  $T(W_1) \subseteq W_1$  implies  $T(v) \in W_1$ . Also  $v \in W_2$  and  $T(W_2) \subseteq W_2$  implies  $T(v) \in W_2$ . Therefore,  $T(v) \in W_1 \cap W_2$ . Since  $v \in W_1 \cap W_2$  is arbitrary we have  $T(W_1 \cap W_2) \subseteq W_1 \cap W_2$ .

## Result

3 of 3

We show that,  $T(W_1 + W_2) \subseteq W_1 + W_2$  and  $T(W_1 \cap W_2) \subseteq W_1 \cap W_2$ .

4. a

Let  $A$  be a  $2 \times 2$  matrix with eigenvectors  $v_1 = (1, 1)^t$  and  $v_2 = (1, 2)^t$  corresponding to the eigenvalues 2 and 3 respectively.

We have to find  $A$ .

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

. Then  $v_1$  is an eigenvectors of  $A$  corresponding to eigenvalue 2 implies

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ i.e. } \begin{bmatrix} a+b \\ c+d \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \text{ i.e. } a+b=2=c+d.$$

Similarly,  $v_2$  is an eigenvectors of  $A$  corresponding to the eigenvalue 3 implies

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ i.e. } \begin{bmatrix} a+2b \\ c+2d \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} \text{ i.e. } a+2b=3 \text{ \& } c+2d=6.$$

Solving these equations we have,  $b = (a + 2b) - (a + b) = 3 - 2 = 1$  and  $d = (c + 2d) - (c + d) = 6 - 2 = 4$  i.e.  $a = 2 - b = 2 - 1 = 1$  and  $c = 2 - d = 2 - 4 = -2$ .

$$\text{So our } A \text{ is } \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}.$$

## Result

3 of 3

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

5. a

We have to find all invariant subspace of the operator  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  defined by

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Certainly, zero vector space is invariant under  $T$ . So let  $V$  be a non-zero  $T$  invariant subspace. Now suppose,  $\begin{bmatrix} a \\ b \end{bmatrix} \in V$  for some  $a, b \in \mathbb{R}$  with  $b \neq 0$ . Then  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ b \end{bmatrix} \in V$ . Now  $\begin{bmatrix} a+b \\ b \end{bmatrix}$  and  $\begin{bmatrix} a \\ b \end{bmatrix}$  is linearly independent : for  $\lambda, \mu \in \mathbb{R}$  with  $\lambda \begin{bmatrix} a+b \\ b \end{bmatrix} + \mu \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  implies that,  $b(\lambda + \mu) = 0$  and  $\lambda(a+b) + \mu a = 0$  i.e.  $\lambda = -\mu$  as  $b \neq 0$ . Hence  $0 = \lambda(a+b) + \mu a = \mu(-a-b+a) = -\mu b$  i.e.  $\mu = 0$  i.e.  $\lambda = 0$ . Now  $\dim(\mathbb{R}^2) = 2$  and  $V$  has two linearly independent vectors, so  $V = \mathbb{R}^2$ .

Now let  $W$  be another non-zero  $T$  invariant subspace such that  $\begin{bmatrix} x \\ y \end{bmatrix} \in W \implies y = 0$ . Then  $\begin{bmatrix} c \\ 0 \end{bmatrix} \in W$  for some  $c \in \mathbb{R} - \{0\}$ . Hence  $\frac{1}{c} \begin{bmatrix} c \\ 0 \end{bmatrix} \in W$  i.e.  $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in W$  i.e.  $W = \left\{ \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} : \alpha \in \mathbb{R} \right\}$ .

Now we have to find all invariant subspaces of the operator  $S : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  defined by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Certainly, zero vector space is invariant. So we have to consider three cases.

**Case 1:**—  $V$  is a invariant subspace under  $S$  and there is a vector  $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in V$  with  $abc \neq 0$ .

Then,  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ 2b \\ 3c \end{bmatrix} \in V$  and  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ 2b \\ 3c \end{bmatrix} = \begin{bmatrix} a \\ 4b \\ 9c \end{bmatrix} \in V$ . Now

let  $\alpha, \beta, \gamma \in \mathbb{R}$  such that,  $\alpha \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \beta \begin{bmatrix} a \\ 2b \\ 3c \end{bmatrix} + \gamma \begin{bmatrix} a \\ 4b \\ 9c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  i.e.  $a(\alpha + \beta + \gamma) = 0$

and  $b(\alpha + 2\beta + 3\gamma) = 0$  and  $c(\alpha + 4\beta + 9\gamma) = 0$ . Since  $abc \neq 0$  we have  $\alpha + \beta + \gamma = 0$  and  $\alpha + 2\beta + 3\gamma = 0$  and  $\alpha + 4\beta + 9\gamma = 0$  and solving these equations we have  $0 = \alpha = \beta = \gamma$ . Hence  $V$  has three linearly independent vectors. Also  $\dim(\mathbb{R}^3) = 3$ . Therefore,  $V = \mathbb{R}^3$ .

**Case 2:**—  $V$  is a invariant subspace under  $S$ . Also let, 1st coordinate of each vector of  $V$  is 0 but  $V$  contains a vector whose 2nd and 3rd coordinates are non-zero.

So let  $\begin{bmatrix} 0 \\ a \\ b \end{bmatrix} \in V$  with  $ab \neq 0$ . Then  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 2a \\ 3b \end{bmatrix} \in V$ . So let

$\alpha, \beta \in \mathbb{R}$  with  $\alpha \begin{bmatrix} 0 \\ a \\ b \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 2a \\ 3b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  i.e.  $a(\alpha + 2\beta) = 0 = b(\alpha + 3\beta)$ .

Since  $ab \neq 0$  we have  $\alpha = 0 = \beta$ . That is  $V$  has two linearly independent vectors. Also  $V$  is a subspace of  $\left\{ \begin{bmatrix} 0 \\ x \\ y \end{bmatrix} : x, y \in \mathbb{R} \right\}$  which has dimension 2.

Therefore,  $V = \left\{ \begin{bmatrix} 0 \\ x \\ y \end{bmatrix} : x, y \in \mathbb{R} \right\}$ .

In similar manner, one can show, if  $V$  is a invariant subspace under  $S$  such that,

2nd (3rd) coordinate of each vector of  $V$  is 0 but  $V$  contains a vector

$$\text{whose 1st and 3rd (2nd) coordinates are non-zero, then } V = \left\{ \begin{bmatrix} x \\ 0 \\ y \end{bmatrix} : x, y \in \mathbb{R} \right\} (= \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} : x, y \in \mathbb{R} \right\}).$$

**Case 3:—**  $V$  is an non-zero invariant subspace under  $S$  such that 1st and 2nd co-ordinates of each vector in  $V$  are zero.

$$\text{So let } \begin{bmatrix} 0 \\ 0 \\ a \end{bmatrix} \in V \text{ with } a \neq 0. \text{ Now } V \text{ is closed under scalar multiplication as it is a vector space, so } V = \left\{ \begin{bmatrix} 0 \\ 0 \\ \lambda a \end{bmatrix} : \lambda \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ x \end{bmatrix} : x \in \mathbb{R} \right\}.$$

In a similar manner, one can show if  $V$  is non-zero  $S$  invariant subspace

such that 2nd and 3rd (1st and 3rd) coordinates of each vector in  $V$  are zero then

$$V = \left\{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} : x \in \mathbb{R} \right\} (= \left\{ \begin{bmatrix} 0 \\ x \\ 0 \end{bmatrix} : x \in \mathbb{R} \right\}).$$

## Result

We showed that, invariant subspace of

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

are zero space,  $\mathbb{R}^2$  and vector space spanned by

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

6. a

Let  $P$  be the real vector space of polynomials  $p(x) = a_0 + a_1x + \dots + a_nx^n$  of degree at most  $n$ .

Let  $D$  denote the derivative  $\frac{d}{dx}$ , considered as linear operator on  $P$ . We have to show  $D$  is a nilpotent operator on  $P$ .

$$\text{Note that } D(b_0 + b_1x + b_2x^2 + b_3x^3 + \dots + b_{m-1}x^{m-1} + a_mx^m) = b_1x + 2b_2x + 3b_3x^2 + \dots + (m-1)b_{m-1}x^{m-2} + mb_mx^{m-1},$$

for  $m \in \{0, 1, 2, \dots, n\}$  and  $b_0, b_1, \dots, b_m \in \mathbb{R}$ . Hence  $D^{n+1} = D \circ D \circ \dots \circ D \{(n+1)\text{-times}\} = 0$  where

Now we find the matrix of  $D$  w.r.t. the basis  $\{1, x, x^2, x^3, \dots, x^n\}$ .

Notice that,  $D(1) = 0, D(x) = 1, D(x^2) = 2x, D(x^3) = 3x^2, \dots, D(x^n) = nx^{n-1}$ . Hence the matrix of  $D$  w.r.t. the basis  $\{1, x, x^2, x^3, \dots, x^n\}$  will be

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & n-1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & n \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$



Now we have to find all invariant subspace of  $D$ .

Let  $V$  be an invariant subspace of positive dimension. Note that, the only subspace of dimension zero is  $\{0\}$  which is certainly invariant.

Let  $f(x) = c_0 + c_1x + c_2x^2 + \dots + c_mx^m \in V$  with  $m \in \{0, 1, 2, \dots, n\}$  and  $c_0, c_1, c_2, \dots, c_m \in \mathbb{R}$  and  $c_m \neq 0$ .

Note that  $D^0(f) := f, D(f), D^2(f), \dots, D^m(f)$  is linearly independent :

for if  $\lambda_0 f + \lambda_1 D(f) + \lambda_2 D^2(f) + \dots + \lambda_m D^m(f) = 0$  for some  $\lambda_0, \dots, \lambda_m \in \mathbb{R}$ , then  $\lambda_0 = 0$  since the coefficient of  $x^m$  in  $D^k(f)$  is 0 when  $1 \leq k \leq m$ .

Hence the linear combination reduces to  $\lambda_1 D(f) + \lambda_2 D^2(f) + \dots + \lambda_m D^m(f) = 0$ . Now,  $\lambda_1 = 0$  as coefficient of  $x^{m-1}$  in  $D^k(f)$  is 0 when  $2 \leq k \leq m$ .

Hence the linear combination reduces to  $\lambda_2 D^2(f) + \dots + \lambda_m D^m(f) = 0$ . Continuing this way, we have  $\lambda_2 = 0, \lambda_3 = 0, \dots, \lambda_m = 0$ .

Let  $q$  be the largest positive integer for which there is a polynomial,

say  $f_q$  in  $V$  of degree  $q$ . Let  $V_q$  be the space of all polynomials with degree at most  $q$ .

Then  $V \subseteq V_q$ . Certainly,  $\dim(V_q) = q + 1$ . Also  $\{f_q, D(f_q), \dots, D^q(f_q)\}$  is a linearly independent subset of  $V$ . Hence  $V = V_q$ .

Therefore, invariant subspace of  $D$  are  $\{0\}$  and  $V_j :=$  set of all polynomials of degree at most  $j$ , where  $j \in \{1, 2, \dots, n\}$ .

## Result

4 of 4

We show that  $D^{n+1} = 0$  and invariant subspace of  $D$  are  $\{0\}$  and  $V_j :=$  set of all polynomials of degree at most  $j$ , where  $j \in \{1, 2, \dots, n\}$ .

## 7. a

- [a] Suppose  $X$  is an eigenvector of the matrix  $A$ . Then  $AX = Y$  is a scalar multiple of  $X$ , so that

$$AX = \lambda X = Y$$

for an eigenvalue  $\lambda$ . Let

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

. Then:

$$(A - \lambda I)x = Ax - \lambda x = 0$$

so that

$$\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (a - \lambda)x_1 + bx_2 \\ cx_1 + (d - \lambda)x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus we arrive at two equations:

$$\begin{aligned} (a - \lambda)x_1 + bx_2 &= 0 \\ cx_1 + (d - \lambda)x_2 &= 0 \end{aligned}$$

Divide both by  $x_2$ , to get equations in  $s$  and  $\lambda$ :

$$\begin{aligned} (a - \lambda) + bs &= 0 \\ c + (d - \lambda)s &= 0 \end{aligned}$$



Eliminating  $\lambda$ , we arrive at the **equation for  $s$** :

$$s = \frac{c + ds}{a + bs}$$

After multiplication with the denominator, this becomes a quadratic equation as it should, since there are two possible eigenvalues, and thus two possible slopes.

- [b)] If an eigenvector  $v$  lies in the third or fourth quadrant, note that  $-v$  will lie in the opposite quadrant and still be an eigenvector. Thus it is enough to show that there is one eigenvector with positive slope and another with negative.

Consider the equation from part a):

$$bs^2 + (a - d)s - c = 0 \iff$$

$$s = \frac{d - a \pm \sqrt{(a - d)^2 + 4bc}}{2b}$$

Assuming that all the  $a, b, c, d$  are positive, we find that

$$d - a + \sqrt{(a - d)^2 + 4bc} > d - a + \sqrt{(a - d)^2} = d - a + |d - a| \geq 0$$

$$d - a - \sqrt{(a - d)^2 + 4bc} < d - a - \sqrt{(a - d)^2} = d - a - |d - a| \leq 0$$

Thus we see that the two solutions have opposite signs, showing that one eigenvector has positive slope and the other negative. This means one can be chosen from the first quadrant and another from the second.

## Result

3 of 3

- Note that  $(A - \lambda I)x = 0$ . From this we get two equations in  $\lambda, x_1, x_2$ . Divide by  $x_1$  to get the slope  $s = x_2/x_1$  into the equations. Then eliminate  $\lambda$ .
- From part a), we solve the equation and show that its solutions have opposite signs. Thus one eigenvector has positive slope and the other negative.

## 8. a

Let  $T$  be a linear operator on a finite dimensional vector space  $V$  over the field  $F$ ,

such that every non-zero vector is an eigenvector. We have to show  $T$  is multiplication by a scalar.

First suppose  $\dim(V) = 1$ . Then there is a non-zero vector  $v \in V$  such that  $V = \{cv : c \in F\}$ . By hypothesis we have  $\lambda \in F$  such that  $T(v) = \lambda v$ , so that  $T(cv) = cT(v) = c(\lambda v) = \lambda(cv)$  i.e.  $T = \lambda I$  where  $I : V \rightarrow V$  is the identity operator. So in this case we are done.

Next suppose,  $\dim(V) \geq 2$ . Then let  $u, w$  be two linearly independent vectors of  $V$ . Now by hypothesis we have  $\alpha, \beta \in F$  such that  $T(u) = \alpha u$  and  $T(w) = \beta w$ . Now note that  $u + w \neq 0$  as  $\{u, w\}$  is a linearly independent set. Hence there is  $\gamma \in F$  such that  $T(u + w) = \gamma(u + w)$ . So that  $\alpha u + \beta w = T(u) + T(w) = T(u + w) = \gamma(u + w)$ . Hence  $\alpha u + \beta w = \gamma u + \gamma w$  i.e.  $(\alpha - \gamma)u = (\gamma - \beta)w$ . Since  $\{u, w\}$  is a linearly independent set we have  $\alpha - \gamma = 0 = \gamma - \beta$  i.e.  $\alpha = \beta$ . What we observe is that for every vector  $w$  which is linearly independent with  $u$  we have  $T(w) = \alpha w$  where  $\alpha \in F$  is such that  $T(u) = \alpha u$ . Now every linearly independent subset can be extended to a basis of  $V$ . So let  $\{v_1, \dots, v_n\}$  be a basis of  $V$  with  $u = v_1$ , then for any  $x \in V$  with representation  $x = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$  with  $c_1, \dots, c_n \in F$  we have  $T(x) = c_1 T(v_1) + c_2 T(v_2) + \dots + c_n T(v_n) = c_1(\alpha v_1) + c_2(\alpha v_2) + \dots + c_n(\alpha v_n) = \alpha(c_1 v_1 + \dots + c_n v_n) = \alpha x$  i.e.  $T = \alpha I$ .

The case when  $\dim(V) = 0$  is trivial as in this case  $V = \{0\}$  so that  $T = 0 = 0I$ , where  $I : V \rightarrow V$  is the identity operator.

## Result

We show that  $T = \delta I$  for some  $\delta \in F$ .

## Section 5

1. a

- [a)] Consider the matrix

$$A = \begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$$

We can write the characteristic polynomial directly since this a  $2 \times 2$  matrix:

$$\begin{aligned} p(t) &= t^2 - (\text{trace } A)t + (\det A) \\ &= t^2 - (-2 + 3)t + (-2 \cdot 3 - (-2) \cdot 2) \\ &= t^2 - t - 2 \end{aligned}$$

Solving  $p(t) = 0$ , we arrive at the two eigenvalues  $\lambda_1 = -1$  and  $\lambda_2 = 2$ . Next, we (try) to find solutions to the systems  $(A - \lambda I)x = 0$ . For  $\lambda = -1$ , we arrive at the system:

$$(A - \lambda I)X = \begin{bmatrix} -1 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which has as solution any vector  $(x_1, x_2)$  with  $x_1 = 2x_2$ . Thus  $(x_1, x_2) = x_2(2, 1)$ , and we see that any multiple of  $(2, 1)$  is an eigenvector. In particular,  $(2, 1)$  itself is an eigenvector.

For  $\lambda = 2$ , we similarly look at the equation

$$(A - \lambda I)X = \begin{bmatrix} -4 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Here we find that  $x_2 = 2x_1$ , and so the eigenvectors are given by  $(2x_1, x_1)$  for arbitrary  $x_1$ . In particular,  $(2, 1)$  is an eigenvector.

- [b)] Consider the matrix

$$A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$

We can write the characteristic polynomial directly since this a  $2 \times 2$  matrix:

$$\begin{aligned} p(t) &= t^2 - (\text{trace} A)t + (\det A) \\ &= t^2 - (1 + 1)t + (1 \cdot 1 - i \cdot (-i)) \\ &= t^2 - 2t \end{aligned}$$

Solving  $p(t) = 0$ , we arrive at the two eigenvalues  $\lambda_1 = 0$  and  $\lambda_2 = 2$ . Next, we (try) to find solutions to the systems  $(A - \lambda I)x = 0$ . For  $\lambda = 0$ , we arrive at the system:

$$(A - \lambda I)X = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which has e.g.  $(1, i)$  as a solution (or any multiple of it). Thus  $(1, i)$  is an eigenvector for  $\lambda = 0$ .

For  $\lambda = 2$ , we similarly look at the equation

$$(A - \lambda I)X = \begin{bmatrix} -1 & i \\ -i & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Here one can see that this is satisfied by the vector  $(-1, i)$ . This is an eigenvector then for  $\lambda = 2$ .

- [c)] Consider the matrix

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

We can write the characteristic polynomial directly since this a  $2 \times 2$  matrix:

$$\begin{aligned} p(t) &= t^2 - (\text{trace} A)t + (\det A) \\ &= t^2 - (\cos \theta + \cos \theta)t + (\cos \theta \cdot \cos \theta - \sin \theta \cdot (-\sin \theta)) \\ &= t^2 - (2 \cos \theta)t + 1 \end{aligned}$$

It is easy to see that this is satisfied by  $x_1 = e^{i\theta}$  and  $x_1 = e^{-i\theta}$ . Next, we (try) to find solutions to the systems  $(A - \lambda I)x = 0$ . We use Euler's formula, which states

$$e^{i\theta} = \cos \theta + i \sin \theta$$

For  $\lambda = e^{i\theta}$ , we arrive at the system:

$$(A - \lambda I)X = \begin{bmatrix} -i \sin \theta & -\sin \theta \\ \sin \theta & -i \sin \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which has e.g.  $(1, -i)$  as a solution (or any multiple of it). Thus  $(1, -i)$  is an eigenvector for  $\lambda = e^{i\theta}$ .

For  $\lambda = e^{-i\theta}$ , we similarly look at the equation

$$(A - \lambda I)X = \begin{bmatrix} i \sin \theta & -\sin \theta \\ \sin \theta & i \sin \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Here one can see that this is satisfied by the vector  $(1, i)$ . This is an eigenvector then for  $\lambda = e^{-i\theta}$ .

## Result

4 of 4

In each of a), b), c) the characteristic polynomial is calculated as  $p(t) = t^2 - (\text{trace} A)t + \det A$ . The quadratic equation  $p(t) = 0$  is solved and then we solve the systems of linear equations  $(A - \lambda I)x = 0$ .

Consider the matrix  $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & a & b \end{pmatrix}$  where  $a, b$  are two complex numbers such that,

$t^3 - 4t - 1$  is the characteristic polynomial of this matrix in indeterminate  $t$ . We have to find  $a$  and  $b$ .

Note that characteristic polynomial of

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & a & b \end{pmatrix}$$

in indeterminate  $t$  is given by

$$\begin{aligned} & \det \begin{pmatrix} t-0 & 1 & 2 \\ 1 & t-1 & 0 \\ 1 & a & t-b \end{pmatrix} \\ &= (t-0)(-1)^{1+1} \begin{vmatrix} t-1 & 0 \\ a & t-b \end{vmatrix} + 1(-1)^{1+2} \begin{vmatrix} 1 & 0 \\ 1 & t-b \end{vmatrix} + 2(-1)^{1+3} \begin{vmatrix} t-1 & 1 \\ 1 & a \end{vmatrix} \\ &= t(t-1)(t-b) - (t-b) + 2a(t-1) - 2 = t^3 - (1+b)t^2 + bt - t + b + 2at - 2a - 2 \\ &= t^3 - (1+b)t^2 - (1-b-2a)t - (2+2a-b). \end{aligned}$$

Now comparing the polynomials  $t^3 - 4t - 1$  and  $t^3 - (1+b)t^2 - (1-b-2a)t - (2+2a-b)$  we have  $1+b=0$  and  $1-b-2a=4$  and  $2+2a-b=1$ , i.e.  $b=-1$  and  $a=-1$ .

## Result

The matrix

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & a & b \end{pmatrix}$$

has characteristic polynomial

$$t^3 - 4t - 1 \text{ in indeterminate } t \text{ if } b = a = -1.$$

3. a

- [a] Suppose a linear operator  $T$  satisfies  $T^r = I$  for some positive integer  $r$ . Suppose  $v$  is an eigenvector of  $T$ . Then

$$\begin{aligned}Tv &= \lambda v \\ T^2 v &= T(\lambda v) = \lambda^2 v \\ &\dots \\ T^r v &= \lambda^r v\end{aligned}$$

Now,  $T^r = I$ , so we also have  $T^r v = Iv = v$ . Since  $v$  is an eigenvector, it is non-zero, thus at least one entry is non-zero. Therefore  $\lambda^r v = v$  implies that

$$\lambda^r = 1$$

We see that an eigenvalue, if there is one, must be an  $r$ -th root of unity. Any of the  $r$ -th roots of unity can be eigenvalues as is exemplified by:

$$T = \begin{bmatrix} e^{\frac{2\pi i}{r}} & 0 & \dots & 0 \\ 0 & e^{\frac{4\pi i}{r}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

- [b] Suppose a linear operator  $T$  satisfies  $(T - 2I)(T - 3I) = 0$ . Suppose  $v$  is an eigenvector of  $T$ . Then

$$\begin{aligned}0 &= (T - 2I)(T - 3I)v = (T - 2I)(\lambda v - 3v) \\ &= (\lambda^2 - 5\lambda + 6)v\end{aligned}$$

Since  $v$  is an eigenvector, it is non-zero, thus at least one entry is non-zero. Therefore  $\lambda^2 - 5\lambda + 6 = 0$  implies that

$$\lambda = 2, 3$$

We see that an eigenvalue, if there is one, must be either  $\lambda = 2, 3$ . Both cases are possible, e.g. by taking  $T = 2I$  or  $T = 3I$ .

## Result

3 of 3

- We show that  $\lambda^r = 1$  so that  $\lambda$  must be an  $r$ -th root of unity.
- We show that  $\lambda$  must be 2 or 3.

4. a



The characteristic polynomial of the matrix in the exercise is calculated as

$$p_n(\lambda) = \det(\lambda I_n - A_n) = \begin{vmatrix} \lambda & -1 & 0 & \dots & 0 \\ -1 & \lambda & -1 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{vmatrix}$$

We will use the multilinearity of determinant. That is the property that the determinant is linear in each column/row. Thus we can write:

$$\begin{aligned} p_n(\lambda) &= \begin{vmatrix} \lambda & -1 & 0 & \dots & 0 \\ -1 & \lambda & -1 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{vmatrix} = \begin{vmatrix} \lambda & 0 & 0 & \dots & 0 \\ -1 & \lambda & -1 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{vmatrix} + \begin{vmatrix} 0 & -1 & 0 & \dots & 0 \\ -1 & \lambda & -1 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{vmatrix} \\ &= \lambda \begin{vmatrix} \lambda & -1 & \dots & 0 \\ -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{vmatrix} - (-1) \begin{vmatrix} -1 & -1 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{vmatrix} \quad (\text{Laplace expansion}) \\ &= \lambda p_{n-1}(\lambda) + (-1) \begin{vmatrix} \lambda & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda \end{vmatrix} \\ &= \lambda p_{n-1}(\lambda) - p_{n-2}(\lambda) \end{aligned}$$

We use the recursion for the characteristic polynomial derived above. For  $n = 1, 2$  we compute it explicitly:

$$\begin{aligned} p_1(\lambda) &= \det(\lambda I_1 - A_1) = \lambda \\ p_2(\lambda) &= \det(\lambda I_2 - A_2) = \lambda^2 - 1 \\ p_3(\lambda) &= \lambda p_2(\lambda) - p_1(\lambda) \\ &= \lambda(\lambda^2 - 1) - \lambda \\ &= \lambda^3 - 2\lambda \\ p_4(\lambda) &= \lambda p_3(\lambda) - p_2(\lambda) \\ &= \lambda(\lambda^3 - 2\lambda) - (\lambda^2 - 1) \\ &= \lambda^4 - 2\lambda^2 - \lambda^2 + 1 \\ &= \lambda^4 - 3\lambda^2 + 1 \\ p_5(\lambda) &= \lambda(\lambda^4 - 3\lambda^2 + 1) - (\lambda^3 - 2\lambda) \\ &= \lambda^5 - 4\lambda^3 + 3\lambda \end{aligned}$$

## Result

3 of 3

Use multilinearity of the determinant in the first row. The recursion turns out to be  $p(\lambda) = \lambda p_{n-1}(\lambda) - p_{n-2}(\lambda)$ .

5. a



Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a  $2 \times 2$  real matrix. Now the characteristic polynomial of  $A$  in indeterminate  $t$  is given by

$$\det \begin{pmatrix} t-a & b \\ c & t-d \end{pmatrix} = (t-a)(t-d) - bc = t^2 - (a+d)t + ad - bc.$$

Now  $A$  has only real eigenvalue

$$\iff t^2 - (a+d)t + ad - bc \in \mathbb{R}[t] \text{ has only real roots}$$

$$\iff \text{discriminate of } t^2 - (a+d)t + ad - bc \text{ is non-negative}$$

$$\iff (-(a+d))^2 - 4 \cdot 1 \cdot (ad - bc) \geq 0$$

$$\iff a^2 + 2ad + d^2 - 4ad + 4bc \geq 0$$

$$\iff (a-d)^2 + 4bc \geq 0$$

Note that when off diagonal entries of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

have same sign then we have  $bc \geq 0$ . Also  $a, d \in \mathbb{R}$  so that,  $(a-d)^2 \geq 0$ . Hence, off diagonal entries have same sign implies  $(a-d)^2 + 4bc \geq 0$  which further implies that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has only real eigenvalues.

## Result

3 of 3

We showed that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has only real eigenvalues if and only if  $(a-d)^2 + 4bc \geq 0$ .

## 6. a

Let  $(v_0, v_1, \dots, v_n)$  be a basis for the vector space  $V$ . Suppose we define a linear operator by defining it on the basis as follows:

$$\begin{aligned} T(v_i) &= v_{i+1}, & 0 \leq i < n \\ T(v_n) &= a_0v_0 + a_1v_1 + \dots + a_nv_n \end{aligned}$$

where the  $a_i$  are scalars.

The matrix of  $T$  is found by writing the coefficients of  $Tv_i$  in the standard basis in the columns. Thus the matrix of the operator is:

$$T = \begin{bmatrix} 0 & 0 & 0 & \dots & a_0 \\ 1 & 0 & 0 & \dots & a_1 \\ 0 & 1 & 0 & \dots & a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{bmatrix}$$

Now consider the characteristic polynomial of the matrix  $T$ , which is computed as  $\det(\lambda I - T)$ . Use Laplace's expansion:

$$\begin{aligned}
 p(\lambda) &= \begin{vmatrix} \lambda & 0 & 0 & \dots & -a_0 \\ -1 & \lambda & 0 & \dots & -a_1 \\ 0 & -1 & \lambda & \dots & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda - a_n \end{vmatrix} \\
 &= \lambda \begin{vmatrix} \lambda & 0 & \dots & -a_1 \\ -1 & \lambda & \dots & -a_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda - a_n \end{vmatrix} + (-1)^n (-a_0) \begin{vmatrix} -1 & \lambda & 0 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -1 \end{vmatrix} \\
 &= \lambda \begin{vmatrix} \lambda & 0 & \dots & -a_1 \\ -1 & \lambda & \dots & -a_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda - a_n \end{vmatrix} + (-1)^n (-a_0) (-1)^n
 \end{aligned}$$

Now, note the self similarity of this expression. We can apply the same logic to the remaining determinant:

$$\begin{aligned}
 &= \lambda \begin{vmatrix} \lambda & 0 & \dots & -a_1 \\ -1 & \lambda & \dots & -a_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda - a_n \end{vmatrix} - a_0 \\
 &= \lambda \left( \lambda \begin{vmatrix} \lambda & \dots & -a_2 \\ \vdots & \vdots & \ddots \\ 0 & \dots & \lambda - a_n \end{vmatrix} - a_1 \right) - a_0 \\
 &= \lambda^2 \begin{vmatrix} \lambda & \dots & -a_2 \\ \vdots & \vdots & \ddots \\ 0 & \dots & \lambda - a_n \end{vmatrix} - a_1 \lambda - a_0 \\
 &\dots \\
 &= \lambda^n (\lambda - a_n) - a_{n-1} \lambda^{n-1} - \dots - a_1 \lambda - a_0 \\
 &= \lambda^{n+1} - a_n \lambda^n - a_{n-1} \lambda^{n-1} - \dots - a_1 \lambda - a_0
 \end{aligned}$$

## Result

3 of 3

The matrix of the operator  $T$  is easy to write down by the definition of a matrix representation of a linear operator.

The characteristic polynomial comes from a determinant which is computed in a recursive way.

7. a

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

. Note that, 1 is the only eigenvalue of  $A$ . Also, if

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

is an eigenvector of  $A$ , then

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 \begin{bmatrix} x \\ y \end{bmatrix}$$

i.e.

$$\begin{bmatrix} x + y \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

i.e.  $y = 0$ . Hence eigenvector of  $A$  are

$$\begin{bmatrix} x \\ 0 \end{bmatrix}$$

, where  $x$  is a non-zero scalar.

Now consider the transpose of  $A$  i.e.

$$A^t = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

. Now notice that 1 is the only eigenvalue of  $A^t$ . So let

$$\begin{bmatrix} a \\ b \end{bmatrix}$$

is an eigenvector of  $A^t$ . Then

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 1 \begin{bmatrix} a \\ b \end{bmatrix}$$

i.e.

$$\begin{bmatrix} a \\ a + b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

i.e.  $a = 0$ . So the eigenvectors of  $A^t$  are

$$\begin{bmatrix} 0 \\ b \end{bmatrix}$$

where  $b$  is a non-zero scalar.

From the above observations, it follows that,  $A$  and  $A^t$  doesn't always have same eigenvector.

Now let  $A$  be a  $n \times n$  matrix over the field  $F$ . Now the characteristic polynomial of  $A$  in indeterminate  $x$  is  $\det(xI_n - A) \in F[x]$ . But notice that, characteristic polynomial of  $A^t$  in indeterminate  $x$  is  $\det(xI_n - A^t) = \det((xI_n - A)^t) = \det(xI_n - A)$ . So the characteristic polynomial of  $A$  and  $A^t$  are same. So if  $\lambda \in F$  is an eigenvalue of  $A$  with algebraic multiplicity  $n$  then,  $\lambda$  is also an eigenvalue of  $A^t$  with algebraic multiplicity  $n$ .

## Result

3 of 3

For a matrix  $A$  eigenvector of  $A$  may not same as eigenvector of  $A$ . But eigenvalues of  $A$  are same as eigenvalues of  $A^t$  (counting multiplicity).

8. a

Consider the  $3 \times 3$  matrix  $A = [a_{ij}]$ . We calculate its characteristic polynomial, in particular the coefficient of the linear term  $t$ . A  $3 \times 3$  determinant can be calculated by e.g. Sarrus' rule.

$$\begin{aligned} p(t) = \det(tI - A) &= \begin{vmatrix} t - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & t - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & t - a_{33} \end{vmatrix} \\ &= (t - a_{11})(t - a_{22})(t - a_{33}) - a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32} + \\ &\quad - (t - a_{11})a_{23}a_{32} - (t - a_{22})a_{13}a_{31} - (t - a_{33})a_{12}a_{21} \\ &= \dots + t(a_{11}a_{22} + a_{22}a_{33} + a_{33}a_{11} - a_{12}a_{21} - a_{23}a_{32}) + \dots \end{aligned}$$

Grouping the terms in the right way, we find that the coefficient is equal to:

$$\begin{aligned} &= a_{11}a_{22} - a_{12}a_{21} + a_{22}a_{33} - a_{23}a_{32} + a_{33}a_{11} - a_{31}a_{13} \\ &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \end{aligned}$$

## Result

2 of 2

Calculate the characteristic polynomial by definition. Collecting the coefficients of  $t$ , it is not hard to check that they are of the given form.

9. a

Consider the linear operator of left multiplication by an  $n \times n$  matrix  $A$  on the space  $F^{n \times n}$  of  $n \times n$  matrices. Calling this operator  $L$ , we are looking for the trace and determinant of the operator

$$L(M) = AM$$

Let  $E_{ij}$  be the matrix with all zeroes except at position  $(i, j)$  where it is 1. The set  $\{E_{ij} \mid 1 \leq i, j \leq n\}$  is the standard basis for  $n \times n$  matrices. Now, consider how  $L$  acts on the subspace spanned by

$$V_r = \text{span}\{E_{1r}, E_{2r}, \dots, E_{nr}\}$$

(these are the matrices who have non-zero entries only in the  $r$ -th column). Consider how the left multiplication acts on these:

$$\begin{aligned} AE_{kr} &= \left[ \sum_{c=1}^n a_{ic}b_{cj} \right] \\ &= \begin{bmatrix} \dots & 0 & a_{1k} & 0 & \dots \\ \dots & 0 & a_{2k} & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \\ \dots & 0 & a_{nk} & 0 & \dots \end{bmatrix} \\ &= \sum_{c=1}^n a_{ck}E_{cr} \in V_r \end{aligned}$$

Here we denote the elements of  $E_{kr}$  with  $b_{ij}$ . Note that  $b_{cj} = 0$  if  $j \neq k$ . The point is that  $AE_{kr}$  has non-zero terms **only** in the  $r$ -th column. Thus we see that  $V_r$  **is an invariant subspace** for left multiplication by  $A$ .

We want to write the operator  $T$  in matrix form with the standard basis. Order the basis as follows:

$$E_{11}, E_{21}, \dots, E_{n1}, E_{12}, E_{22}, \dots, E_{2n}, \dots, E_{n1}, \dots, E_{nn}$$

Thus the first  $n$  columns will be determined by the action of  $A$  on  $E_{1r}, E_{2r}, \dots, E_{nr}$  and so on. As seen above,  $AV_r \subseteq V_r$ , thus we in fact get a block matrix of the form:

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_n \end{bmatrix}$$

where  $A_i$  is the matrix of the operator acting  $A$  acting on  $V_r$ . As shown above,

$$AE_{kr} = A_i E_{kr} = \sum_{c=1}^n a_{ck} E_{cr}$$

We see that this expression does not depend on  $i$ . Thus **all** the matrices  $A_i$  are one and the same. It is easy to check that in fact

$$A_i = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = A$$

Finally, all of this means that the operator  $L$  has as matrix the block matrix with  $A$  on the diagonal

$$L = \begin{bmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A \end{bmatrix}$$

The trace and determinant are computed as

$$\begin{aligned} \text{trace } L &= \text{trace } A + \text{trace } A + \dots + \text{trace } A \\ &= n \cdot \text{trace } A \\ \det L &= \det A \cdot \det A \dots \det A \\ &= (\det A)^n \end{aligned}$$

## Result

5 of 5

Consider the standard basis for  $n \times n$  matrices given by  $E_{ij} = [\delta_{ij}]$ . One can check directly that  $AE_{kr} = \sum_{c=1}^n a_{ck} E_{cr}$ . Thus grouping the basis matrices appropriately, we find that the matrix for left multiplication by  $A$  (in the standard basis) can be written as a diagonal block matrix. The trace and determinant are then easily related to the trace and determinant of  $A$ .

10. a

Consider the linear operator of multiplication by  $n \times n$  matrices  $A = [a_{ij}]$ ,  $B = [b_{ij}]$  on the space  $F^{n \times n}$  of  $n \times n$  matrices given by

$$T(M) = AMB$$

In the previous exercise we calculated the determinant of the operator  $L$  defined by  $L(M) = AM$ . Consider now the operator of right multiplication

$$R(M) = MB$$

As in the previous exercise, let  $E_{ij}$  be the matrix with all zeroes except at position  $(i, j)$  where it is 1. The set  $\{E_{ij} \mid 1 \leq i, j \leq n\}$  is the standard basis for  $n \times n$  matrices. Similar to the previous calculation, we can consider now

$$W_r = \text{span}\{E_{r1}, E_{r2}, \dots, E_{rn}\}$$

We would find that

$$\begin{aligned} R(E_{rk}) &= E_{rk}B \\ &= [b_{ij}\delta_{ki}] \\ &= \sum_{c=1}^n b_{kc}E_{rc} \end{aligned}$$

Again, just as in the previous example, we would find that  $R$  is a diagonal block matrix

$$B = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_n \end{bmatrix}$$

Again,  $B_i E_{rk} = B E_{rk}$  and so all the  $B_i$  are the same and given by

$$B_i = \begin{bmatrix} b_{11} & b_{21} & \dots & b_{n1} \\ b_{12} & b_{22} & \dots & b_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1n} & b_{n2} & \dots & b_{nn} \end{bmatrix} = B^t$$

This means that

$$\det R = (\det B^t)^n = (\det B)^n$$

Notice that our operator  $T$  can be written as a composition of left multiplication by  $A$  and right multiplication by  $B$  i.e.

$$T(M) = LR(M) = AMB$$

The determinant of this operator is

$$\det(T) = \det(LR) = \det(L) \det(R) = (\det A)^n (\det B)^n$$



The trace is not multiplicative so the same proof doesn't apply. Consider how  $T(M) = AMB$  acts on a basis vector

$$\begin{aligned} T(E_{kr}) &= AE_{kr}B = \left[ \sum_{c=1}^n a_{ic} \delta_{cj} \right] B \\ &= [a_{ij} \delta_{kj}] B \\ &= \left[ \sum_{c=1}^n a_{ic} \delta_{kc} b_{cj} \right] \\ &= [a_{ik} b_{rj}] \\ &= \dots + a_{kk} b_{rr} E_{kr} + \dots \end{aligned}$$

Thus, writing out  $T$  in the standard basis, we will find that the diagonal entries have coefficients

$$a_{kk} b_{rr}, \quad 1 \leq k, r \leq n$$

The trace is the sum of the main diagonal elements, so

$$\begin{aligned} \text{trace } T &= \sum_{k=1}^n \sum_{r=1}^n a_{kk} b_{rr} \\ &= \left( \sum_{k=1}^n a_{kk} \right) \left( \sum_{r=1}^n b_{rr} \right) \\ &= (\text{trace } A)(\text{trace } B) \end{aligned}$$

## Result

5 of 5

Using the previous exercise and an analogous statement for right multiplication, we prove that the operator has determinant  $(\det A)^n (\det B)^n$ . The trace is calculated by considering an explicit basis and turns out to be the product of the traces of  $A$  and  $B$ .

## Section 6

1. a

Let  $A$  be an  $n \times n$  matrix whose characteristic polynomial factors into linear factors:

$p(t) = (t - \lambda_1) \dots (t - \lambda_n)$ . By the textbook lemma, this means that there is a set of eigenvectors which form a basis in which  $A$  becomes diagonal. This means that there is a similarity matrix such that

$$P^{-1}AP = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Now recall that  $\text{trace } AB = \text{trace } BA$ . Thus

$$\begin{aligned} \text{trace } A &= \text{trace } (AP)P^{-1} \\ &= \text{trace } P^{-1}(AP) \\ &= \text{trace } \Lambda \\ &= \lambda_1 + \lambda_2 + \dots + \lambda_n \end{aligned}$$

For the determinant it's even easier:

$$\begin{aligned} \det A &= (\det P)^{-1} \det A(\det P) \\ &= \det P^{-1}AP \\ &= \det \Lambda \\ &= \lambda_1 \lambda_2 \dots \lambda_n \end{aligned}$$

## Result

2 of 2

Since the characteristic polynomial factors completely, the matrix  $A$  is diagonalizable. The resulting diagonal  $\Lambda = P^{-1}AP$  matrix has the eigenvalues as entries. It is easy to see that  $A$  and  $\Lambda$  have the same trace and determinant.

## 2. a

- [a)] Suppose a complex  $n \times n$  matrix  $A$  has distinct eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  with eigenvectors  $v_1, \dots, v_n$ . Eigenvectors coming from distinct eigenvalues must form a linearly independent set. Since there are  $n$  of them, we see that they form a basis for  $\mathbb{C}^n$ . An arbitrary vector represented in this basis is of the form

$$v = a_1 v_1 + a_2 v_2 + \dots + a_n v_n$$

Suppose  $v$  is an arbitrary eigenvector i.e.  $Av = kv$  for some  $k \in \mathbb{C}$ . All the eigenvalues are given by the  $\lambda_i$  thus  $k = \lambda_i$ .

Suppose then that  $Av = \lambda_i v = \sum_{j=1}^n \lambda_i a_j v_j$ . But

$$Av = A\left(\sum_{j=1}^n a_j v_j\right) = \sum_{j=1}^n \lambda_j a_j v_j$$

This leads to two different representations of the same vector:

$$\begin{aligned} Av &= \sum_{j=1}^n \lambda_i a_j v_j = \sum_{j=1}^n \lambda_j a_j v_j \implies \\ \sum_{j=1}^n (\lambda_i - \lambda_j) a_j v_j &= 0 \end{aligned}$$

Since the  $v_j$  are linearly independent, we must have that the coefficients are all zero. But since  $\lambda_i$  are distinct, we find that all the coefficients  $a_j$  must be zero (except for  $j = i$ ). We see that  $v$  must be of the form

$$v = a_i v_i$$

Thus if  $v$  is eigenvector, it must belong to one of the given eigenvalues and must in fact be a multiple of the corresponding  $v_i$ .

- [b)] Suppose the conditions above hold. Write down the matrix  $[B]$  which has as columns the eigenvectors  $v_1, \dots, v_n$  (in that order). This is a change of basis matrix for the eigenvector basis. We know that

$$\Lambda = [B]^{-1}A[B]$$

where  $\Lambda$  is the diagonal matrix which has the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  (in that order). Thus  $\Lambda$  is easy to write down as is  $[B]$ .  $[B]^{-1}$  exists and is calculated in the usual way. Thus

$$A = [B]\Lambda[B]^{-1}$$

## Result

3 of 3

a) Note that the eigenvectors form a basis.  $\lambda_1, \dots, \lambda_n$  are all the eigenvalues, so assume  $Av = \lambda_i v$ . Write  $v$  down in the eigenvectors basis. Equating coefficients, one will find that the only possibility is for  $v$  to be a multiple of  $v_i$ .

b) We know that  $\Lambda = [B]^{-1}A[B]$  where  $\Lambda$  is the diagonal matrix with eigenvalues on its diagonal, and  $[B]$  is a change of basis matrix for the eigenvectors (which do form a basis). From this, it is easy to recover  $A$ .

## 3. a

Let  $T$  be a linear operator on a finite dimensional vector space  $V$  and  $\lambda$  be an eigenvalue of  $T$

such that  $v_1, v_2$  are two linearly independent eigenvectors of  $T$  corresponding to the eigenvalue  $\lambda$ .

We have to show multiplicity of  $\lambda$  in characteristic polynomial is at least 2.

Since  $\{v_1, v_2\}$  is a linearly independent subset of  $V$  we can extend this subset to a basis of  $V$ , say  $\mathcal{B} = \{v_1, v_2, v_3, \dots, v_n\}$  is a basis of  $V$ . Note that,  $T(v_k) = \lambda v_k$  for  $k = 1, 2$ . Now the matrix of  $T$  w.r.t.  $\mathcal{B}$  is of the following form :---

$$M = \begin{bmatrix} \lambda I_2 & B \\ 0 & C \end{bmatrix},$$

where  $I_2$  denotes the  $2 \times 2$  identity matrix and  $B$  is a  $2 \times (n-2)$  matrix and  $C$  is a  $(n-2) \times (n-2)$  matrix and  $0$  is the  $(n-2) \times 2$  zero matrix. Hence the characteristic polynomial  $\chi_M(x)$  of  $M$  in indeterminate  $x$  is  $(x - \lambda)^2 \chi_C(x)$ , where  $\chi_C(x)$  denotes the characteristic polynomial of  $C$  in indeterminate  $x$ . That's eigenvalue  $\lambda$  of the operator  $T$  has multiplicity at least 2 in the characteristic polynomial of the matrix  $M$ . So we are done.

## Result

3 of 3

We show that multiplicity of  $\lambda$  in characteristic polynomial of  $T$  is at least 2.

## 4. a

Note that the characteristic polynomial of

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

is  $x^2 - 4x + 3 = (x - 3)(x - 1)$ . So the eigenvalues are 3 and 1.  
Also,

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

So let

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

then

$$\begin{aligned} P^{-1}AP &= P^{-1}A[P_{*1} : P_{*2}] = P^{-1}[AP_{*1} : AP_{*2}] \\ &= P^{-1}[3P_{*1} : 1P_{*2}] = P^{-1}P \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Here  $P_{*1}, P_{*2}$  denote 1st and 2nd columns of  $P$  respectively.

Now notice that  $(P^{-1}AP)^{30} = P^{-1}A^{30}P$ . So that,

$$A^{30} = P \left( (P^{-1}AP)^{30} \right) P^{-1} = P \begin{bmatrix} 3^{30} & 0 \\ 0 & 1^{30} \end{bmatrix} P^{-1}.$$

Now note that

$$P^{-1} = \frac{1}{-2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$$

, so

$$A^{30} = \frac{1}{-2} \begin{bmatrix} 3^{30} & 1 \\ 3^{30} & -1 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3^{30} + 1 & 3^{30} - 1 \\ 3^{30} - 1 & 3^{30} + 1 \end{bmatrix}.$$

## Result

We show that

$$A^{30} = \frac{1}{2} \begin{bmatrix} 3^{30} + 1 & 3^{30} - 1 \\ 3^{30} - 1 & 3^{30} + 1 \end{bmatrix}$$

5. a

- [a] Since there is no obvious way to write down such a matrix, we will compute the eigenvectors. If the matrix is diagonalizable, then the eigenvectors will form a basis. The change of basis matrix will give the required similarity matrix  $P$ .

First we compute the characteristic polynomial:

$$\begin{aligned} p(t) &= \det(tI - A) = \begin{vmatrix} t-1 & -i \\ i & t-1 \end{vmatrix} \\ &= (t-1)^2 - i(-i) \\ &= t^2 - 2t \end{aligned}$$

It has the zeroes  $t_{1,2} = 0, 2$  and these are the eigenvalues  $\lambda_1, \lambda_2$ . Now we find the eigenvectors by computing  $\ker A - \lambda I$ . For  $\lambda = 0$ , one has

$$(A - \lambda I)v = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This is realized for  $v_1 = (x_1, x_2) = (1, i)$  (and any multiple of this).

Consider now the other eigenvalue  $\lambda = 2$ . One has

$$(A - \lambda I)v = \begin{bmatrix} -1 & i \\ -i & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

One finds that this is satisfied by  $v_2 = (1, -i)$ . This gives another eigenvector.

Write the eigenvectors as columns to get the change of basis matrix

$$P = [B] = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

From theory, we know that  $P^{-1}AP$  is diagonal, thus  $P$  is the required matrix.

- [b] First we compute the characteristic polynomial:

$$\begin{aligned} p(t) &= \det(tI - A) = \begin{vmatrix} t & 0 & -1 \\ -1 & t & 0 \\ 0 & -1 & t \end{vmatrix} \\ &= t^3 - 1 \end{aligned}$$

It has the zeroes  $t_{1,2,3} = 1, \omega, \omega^2$  ( $\omega = \frac{-1-i\sqrt{3}}{2}$ ) and these are the eigenvalues  $\lambda_1, \lambda_2, \lambda_3$ . Now we find the eigenvectors by computing  $\ker A - \lambda I$ . For  $\lambda = 1$ , one has

$$(A - \lambda I)v = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

It is easily guessed that this is satisfied by  $(1, 1, 1)$ . Since all the eigenvalues are distinct, there can only be one eigenvector for each eigenvalue (up to multiples).

Consider now the eigenvalue  $\lambda = \omega$ . One has

$$(A - \lambda I)v = \begin{bmatrix} \omega & 0 & 1 \\ 1 & \omega & 0 \\ 0 & 1 & \omega \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$



By summing the rows, we would find that  $x_1 + x_2 + x_3 = 0$ . From this one can infer that a solution is given by

$$(x_1, x_2, x_3) = (1, \omega^2, -\omega).$$

The same reasoning works for  $\lambda = \omega^2$ , where one gets the eigenvector  $(x_1, x_2, x_3) = (1, \omega, -\omega^2)$ .

Write the eigenvectors as columns of the change of basis matrix:

$$P = [B] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \omega^2 & -\omega \\ 0 & \omega & -\omega^2 \end{bmatrix}$$

From theory, we know that  $P^{-1}AP$  is diagonal, thus  $P$  is the required matrix.

- [c]

First we compute the characteristic polynomial:

$$\begin{aligned} p(t) &= \det(tI - A) = \begin{vmatrix} t - \cos \theta & \sin \theta \\ -\sin \theta & t - \cos \theta \end{vmatrix} \\ &= (t - \cos \theta)^2 - \sin \theta(-\sin \theta) \\ &= t^2 - (2 \cos \theta)t + 1 \end{aligned}$$

It has the zeroes  $t_{1,2} = e^{i\theta}, e^{-i\theta}$  and these are the eigenvalues  $\lambda_1, \lambda_2$ . Now we find the eigenvectors by computing  $\ker A - \lambda I$ . For  $\lambda = e^{i\theta}$ , one has

$$(A - \lambda I)v = \begin{bmatrix} -i \sin \theta & \sin \theta \\ -\sin \theta & -i \sin \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This is realized for  $v_1 = (x_1, x_2) = (1, i)$ .

Consider now the other eigenvalue  $\lambda = e^{-i\theta}$ . One has

$$(A - \lambda I)v = \begin{bmatrix} i \sin \theta & \sin \theta \\ -\sin \theta & i \sin \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

One finds that this is satisfied by  $v_2 = (1, -i)$ . This gives another eigenvector.

Write the eigenvectors as columns to get the change of basis matrix

$$P = [B] = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

From theory, we know that  $P^{-1}AP$  is diagonal, thus  $P$  is the required matrix.

## Result

4 of 4

In each of a),b),c), one first computes the characteristic polynomial. The zeroes of this polynomial are the eigenvalues. Finally, solving the system  $(A - \lambda I)v = 0$  for each eigenvalue, gives an eigenvector.  $P$  is given as a change of basis matrix from the standard basis to the eigenvector basis.

6. a



Let  $A$  be a  $n \times n$  matrix over a field  $F$  such that  $A$  is diagonalizable over  $F$  i.e. there is an invertible  $n \times n$  matrix  $P$  over  $F$  such that  $PAP^{-1} = D$ , where  $D$  is a diagonal matrix over  $F$ . We show that there is a  $n \times n$  matrix  $Q$  over  $F$  such that  $\det(Q) = 1$  and  $QAQ^{-1} = D$ .

Note that,  $\det(P) \neq 0$  as  $PP^{-1} = I$  implies  $\det(P)\det(P^{-1}) = \det(I) = 1$ . So let  $Q = \frac{1}{\det(P)}P$ , then  $\det(Q) = 1$  and  $Q^{-1} = \det(P)P^{-1}$ . Now,

$$\begin{aligned} QAQ^{-1} &= \left( \frac{1}{\det(P)}P \right) A \left( \det(P)P^{-1} \right) \\ &= \left( \frac{1}{\det(P)}\det(P) \right) PAP^{-1} = 1D = D. \end{aligned}$$

## Result

2 of 2

$$PAP^{-1} = D \text{ implies } \left( \frac{1}{\det(P)}P \right) A \left( \det(P)P^{-1} \right) = \left( \frac{1}{\det(P)}\det(P) \right) PAP^{-1} = 1D = D.$$

7. a

Let  $A$  and  $B$  be two  $n \times n$  matrices over the field  $F$  such that  $A$  is non-singular. Then  $BA = A^{-1}(AB)A = A^{-1}(AB)(A^{-1})^{-1}$  where,  $A^{-1}$  denotes the inverse of  $A$ . Therefore,  $AB$  is similar to  $BA$ .

## Result

2 of 2

$$A \text{ is non-singular implies that } BA = A^{-1}(AB)A = A^{-1}(AB)(A^{-1})^{-1}.$$

8. a

Method 1.

To prove that  $T$  is nilpotent if and only if there is a basis of  $V$  such that the matrix  $T$  is upper triangular with diagonal entries zero,

Suppose  $T$  be the linear operator and is nilpotent for some positive integer  $k$ .

Then,

$$T^k = 0$$

Then, show that there is a basis of  $V$  such that the matrix  $T$  is upper triangular with diagonal entries zero.

Suppose  $\lambda$  be the eigenvalue of the matrix  $T$ ,

Then,

$$\lambda^k = 0$$

This implies that,

$$\lambda = 0$$

This implies that  $0$  is the only eigenvalue of the nilpotent matrix  $T$ .

Then, corresponding to the each eigenvalue  $0$ , there is an eigenvector  $v_i$ .

Then, the eigenvector  $(v_i)$  can be extended to form the basis  $B = (v_1, v_2, \dots, v_n)$  for the vector space  $V$ .

Then, the matrix  $T$  will have the form

$$T = \begin{bmatrix} \lambda & a_{12} & a_{13} & a_{14} & \cdots & \cdots & a_{1n} \\ & \lambda & a_{23} & a_{24} & \cdots & \cdots & a_{2n} \\ & & \lambda & a_{34} & \cdots & \cdots & a_{3n} \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \\ & & & & & \lambda & a_{(n-1)n} \\ & & & & & & \lambda \end{bmatrix}$$

But, all the eigenvalue of the matrix  $T$  are zero.

That is,

$$\lambda = 0$$

Then,

$$T = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} & \cdots & \cdots & a_{1n} \\ & 0 & a_{23} & a_{24} & \cdots & \cdots & a_{2n} \\ & & 0 & a_{34} & \cdots & \cdots & a_{3n} \\ & & & \ddots & \ddots & \ddots & \vdots \\ & & & & \ddots & \ddots & \vdots \\ & & & & & 0 & a_{(n-1)(n-1)} \\ & & & & & & 0 \end{bmatrix}$$

Hence, it concludes that if the matrix  $T$  is nilpotent then there is a basis of  $V$  such that the matrix  $T$  is upper triangular with diagonal entries zero.

Converse part,

Suppose  $B$  a basis of the vector space  $V$  such that the matrix  $T$  is upper triangular with diagonal entries zero.

Then,

$$T = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} & \cdots & \cdots & a_{1n} \\ & 0 & a_{23} & a_{24} & \cdots & \cdots & a_{2n} \\ & & 0 & a_{34} & \cdots & \cdots & a_{3n} \\ & & & \ddots & \ddots & \ddots & \vdots \\ & & & & \ddots & \ddots & \vdots \\ & & & & & 0 & a_{(n-1)(n-1)} \\ & & & & & & 0 \end{bmatrix}$$

Since, the above matrix  $T$  is the strict triangular matrix.

Then, the matrix  $T$  must vanish at  $n^{\text{th}}$  power.

Then,

$$T^n = 0$$

This implies that the matrix  $T$  is nilpotent.

Therefore, the matrix  $T$  is nilpotent if  $B$  is a basis of the vector space  $V$  such that the matrix  $T$  is upper triangular with diagonal entries zero.

Hence, **it is proved that  $T$  is nilpotent if and only if there is a basis of  $V$  such that the matrix  $T$  is upper triangular with diagonal entries zero.**

## Method 2.

Let  $T$  be a nilpotent operator over a finite dimensional vector space  $V$  i.e.  $T^k = 0$  for some positive integer  $k$ .

We have to show there is a basis of  $V$  for which matrix of  $T$  is upper triangular, with diagonal entries are 0.

**Proof :—** The result is true if  $T$  is zero operator. So we may assume,  $T$  is a non-zero operator.

We prove this by induction on the dimension of  $V$ . Clearly, this result holds when  $\dim(V) = 1$ . So assume,  $\dim(V) > 1$ .

Also let the result holds for each vector space with dimension strictly less than dimension of  $V$ .

Null space of  $T$  has a non-zero element : Let  $n$  be the smallest positive integer for which  $T^n = 0$ . Then  $T^{n-1} \neq 0$ .

So there is  $v \in V$  with  $T^{n-1}(v) \neq 0$ . But  $T^n = 0$  implies  $T(T^{n-1}(v)) = 0$ .

Therefore, dimension of null space of  $T$  is at least 1. So by dimension formula we have  $\dim(\text{range}(T)) < \dim(V)$ .

Consider the operator  $S = T|_{\text{range}(T)} : \text{range}(T) \rightarrow \text{range}(T)$ . By our induction hypothesis,

we have a basis  $\{u_1, u_2, \dots, u_m\}$  of  $\text{range}(T)$  for which matrix of  $S$  is upper triangular.

Extend the basis  $\{u_1, \dots, u_m\}$  of  $\text{range}(T)$  to a basis of  $V$ , say  $\{u_1, \dots, u_m, u_{m+1}, \dots, u_l\}$  is a basis for  $V$ .

Now notice that for each  $p \in \{m+1, \dots, l\}$  we have  $T(u_p) \in \text{range}(T) = \text{span}\{u_1, u_2, \dots, u_m\} \subseteq \{u_1, \dots, u_p\}$ .

Hence the matrix  $M$  of  $T$  w.r.t.  $\{u_1, \dots, u_l\}$  is upper triangular.

Now characteristic polynomial of  $T$  in indeterminate  $t$  is given by  $\det(tI_l - M)$ . But notice that, as  $M$  is upper triangular,

so is  $tI_l - M$ . Also determinant of an upper triangular matrix is nothing but product of all diagonal entries.

Hence  $\det(tI_l - M) = (t - M_{11}) \dots (t - M_{ll})$ , where  $M_{pp}$  denotes the  $p$ -th diagonal entry of  $M$  for  $p = 1, \dots, l$ .

Hence eigenvalue of  $M$  as well as  $T$  are  $M_{11}, \dots, M_{ll}$ . Now for each  $p \in \{1, \dots, l\}$  we have non-zero vector  $v_p \in V$

such that  $T(v_p) = M_{pp}v_p$  which implies,  $T^2(v_p) = M_{pp}^2v_p, T^3(v_p) = M_{pp}^3v_p, \dots, T^k(v_p) = M_{pp}^k v_p$ . But by hypothesis  $T^k = 0$ .

Hence  $M_{pp}^k = 0$  as  $v_p \neq 0$  i.e.  $M_{pp} = 0$  for each  $p \in \{1, \dots, l\}$ .

Now let  $T$  be a linear operator on a finite dimensional vector space  $V$  such that with respect to some basis of  $V$  the matrix of  $T$  is upper triangular with diagonal entries are zero. We have to show there is a positive integer  $d$  such that  $T^d = 0$ .

So let  $\{v_1, \dots, v_n\}$  be a basis for  $V$  with respect to which matrix  $A = [a_{ij}]_{i,j=1}^n$  of  $T$  is upper triangular, with diagonal entries are zero. Therefore,  $a_{ij} = 0$  if  $i \geq j$  and  $i, j \in \{1, \dots, n\}$ . Also, we have  $T(v_j) = \sum_{i=1}^n a_{ij}v_i$  for each  $j \in \{1, \dots, n\}$ .

Hence

$$\begin{aligned} T(v_1) &= 0 \implies T^k(v_1) = 0 \text{ if } k \geq 1, \\ T(v_2) &= a_{12}v_1 \implies T^k(v_2) = 0 \text{ if } k \geq 2, \\ T(v_3) &= a_{13}v_1 + a_{23}v_2 \implies T^k(v_3) = 0 \text{ if } k \geq 3, \\ T(v_4) &= a_{14}v_1 + a_{24}v_2 + a_{34}v_3 \implies T^k(v_4) = 0 \text{ if } k \geq 4, \\ &\vdots \\ T(v_n) &= a_{1n}v_1 + \dots + a_{(n-1)n}v_{n-1} \implies T^k(v_n) = 0 \text{ if } k \geq n. \end{aligned}$$

Using these we have  $T^n = 0$ . Hence  $T$  is nilpotent.

## Result

We do it by induction on dimension of  $V$ .

9. a

Let  $A$  be a real  $2 \times 2$  matrix and suppose  $A^2 = I$ . This amount to the matrix equality:

$$A^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus we arrive at the set of equations:

$$\begin{aligned} a^2 + bc &= 1 \\ b(a+d) &= 0 \\ c(a+d) &= 0 \\ d^2 + bc &= 1 \end{aligned}$$

We distinguish two cases

- If  $a + d \neq 0$ , then  $b = c = 0$ . The first and fourth equation reduce to  $a^2 = d^2 = 1$ . Since  $a + d \neq 0$ , we find that  $a = d$  and this gives the possible matrices

$$A = \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix} = \pm I$$

- Suppose  $a + d = 0$ . Consider the characteristic polynomial. For a  $2 \times 2$  matrix this is

$$p(t) = t^2 - (a + d)t + ad - bc$$

Since  $a + d = 0$ , we have

$$p(t) = t^2 - a^2 - bc = t^2 - 1$$

in view of the first equality above. Thus the matrix **does** have **real** eigenvalues. If they are distinct, then the matrix can be diagonalized. Assume that the matrix only has one eigenvalue. Then it is similar to an upper triangular matrix of the form

$$B = \begin{bmatrix} \pm 1 & x \\ 0 & \pm 1 \end{bmatrix}$$

But the trace of this matrix is non-zero, whereas similar matrices must have the same trace. (The trace of  $A$  is  $a + d = 0$ .)

Now we look at the case when the eigenvalues are distinct. This means there is a basis with eigenvectors  $v_1, v_2$ , and that the matrix  $A$  is diagonalizable. Write

$$P = [v_1 \quad v_2]$$

for the change of basis matrix. Then

$$A = P^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} P$$

We see that either  $A = \pm I$  or  $A$  has distinct eigenvalues and

$$A = P^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} P$$

where  $P$  is the change of basis matrix for the eigenvector basis.

Left multiplication by  $I$  doesn't change the coordinates. Left multiplication by  $-I$  rotates the coordinates  $180^\circ$  around the origin.

Finally,

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

is just a reflection around the  $y$ -axis (if we reverse the order of the eigenvalues we'll have a reflection around the  $x$ -axis but this is not relevant).

Therefore,

$$A = P^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} P$$

changes the coordinate basis, does a reflection in the new coordinates, and then turns the coordinates back.



## Result

4 of 4

Write down a general  $2 \times 2$  matrix  $A$ . If the trace is non-zero, then  $A = \pm I$  and this is the identity or rotation by  $180^\circ$  around the origin.

If the trace is 0, then the characteristic polynomial of  $A$  is  $t^2 - 1$ . The trace of  $A$  being 0, the eigenvalues of  $A$  must sum to zero, thus the eigenvalues are  $1, -1$ . This shows that

$$A = P^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} P$$

The matrix in-between is just a reflection around the  $y$ -axis. Thus  $A$  changes the basis, reflects, and changes back.

10. a

( $\Rightarrow$ ) Suppose  $A, D$  are diagonalizable matrices. Write

$$M = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$$

Let  $e'_1, \dots, e'_n$  and  $f'_1, \dots, f'_m$  be eigenvectors for  $A, D$ .

Extend the  $e'_i$  to  $e_i$  by adding  $m$  by concatenating  $m$  zeroes. Extend  $f'_i$  to  $f_i$  by concatenating  $n$  zeroes from the front. Then

$$\begin{aligned} M e_i &= \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} e'_i \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} A e'_i \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} k e'_i \\ 0 \end{bmatrix} \\ &= k e_i \end{aligned}$$

for some scalar  $k$ . The same applies to the  $f_i$ . It is obvious that  $\{e_1, \dots, e_n, f_1, \dots, f_m\}$  is a linearly independent set. Thus we have constructed a basis of eigenvectors. We conclude that in this basis the matrix is diagonal.

( $\Leftarrow$ ) Now suppose a block matrix  $M$  is diagonalizable. Suppose  $A$  was an  $n \times n$  matrix and  $D$  an  $m \times m$  matrix. Since  $M$  is diagonalizable,  $M$  must have a basis of eigenvectors  $v_1, v_2, \dots, v_n, v_{n+1}, \dots, v_{n+m}$ . Write each of these as

$$v_i = e_i + f_i$$

where  $e_i$  has zeroes in the last  $m$  entries and  $f_i$  has zeroes in the first  $n$  entries. Consider the dimension of subspaces

$$\begin{aligned} V &= \text{span}\{e_i \mid i = 1, \dots, n+m\} \\ W &= \text{span}\{f_i \mid i = 1, \dots, n+m\} \end{aligned}$$

Since  $e_i$  has **only** the first  $n$  entries non-zero, we see that  $\dim V \leq n$ . Similarly,  $\dim W \leq m$ . Now,  $V \cap W = \{0\}$  and

$$\begin{aligned} \dim(V + W) &= \dim \text{span}\{e_i, f_i \mid i = 1, \dots, n+m\} \\ &= \dim \text{span}\{v_i \mid i = 1, \dots, n+m\} \\ &= n + m \end{aligned}$$

Finally, we arrive at the inequality:

$$\begin{aligned} n + m &= \dim(V + W) \\ &= \dim V + \dim W - \dim(V \cap W) \\ &\leq n + m \end{aligned}$$



We see that the inequality must in fact be an equality and this happens if and only if  $\dim V = n$  and  $\dim W = m$ . Reindexing if necessary, we see that  $V$  is generated by exactly  $n$  of the  $e_i$  and similarly for  $W$ , so:

$$\begin{aligned} V &= \text{span}\{e_k \mid k = 1, \dots, m\} \\ W &= \text{span}\{f_k \mid k = 1, \dots, m\} \end{aligned}$$

Since  $V + W$  is the whole vector space, we see that the eigenvector basis  $v_1, \dots, v_{n+m}$  can be replaced by  $e_1, \dots, e_n, f_1, \dots, f_m$ .

Denote by  $e'_i$  the vector  $e_i$  from which the last  $m$  terms have been truncated. One can then write:

$$\begin{aligned} \begin{bmatrix} Ae'_i \\ 0 \end{bmatrix} &= \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e'_i \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} e'_i \\ 0 \end{bmatrix} \end{aligned}$$

Note that we added the  $D$  with no effect since it will just multiply with zeroes. Thus this vector is expressible only in the  $e_i$  (i.e.

$$\begin{bmatrix} Ae'_i \\ 0 \end{bmatrix} \in V$$

).

Next, we bring this into connection with the eigenvector  $v_i$ :

$$\begin{aligned} \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} e'_i \\ 0 \end{bmatrix} &= Me_i \\ &= M(v_i - f_i) \\ &= kv_i - Mf_i \\ &= k(e_i + f_i) - Mf_i \in V \end{aligned}$$

Note that  $Mf_i \in W$  and since  $V$  and  $W$  are disjoint, we can be sure that there are no  $e_i$  in the expression. The end result must be in  $V$ , so we conclude that  $kf_i = Mf_i$  and finally that

$$\begin{bmatrix} Ae'_i \\ 0 \end{bmatrix} = ke_i = \begin{bmatrix} ke'_i \\ 0 \end{bmatrix}$$

Thus  $e'_i$  is an eigenvector for  $A$ . Obviously,  $e'_1, \dots, e'_n$  is a basis. Similarly, we would find that by truncating the first  $n$  terms of the  $f_i$ , would give us eigenvectors for  $D$ . Thus we conclude that both  $A$  and  $D$  have a basis in eigenvectors, therefore, they are diagonalizable.

## Result

4 of 4

If  $A, D$  are diagonalizable, then the eigenvector bases for  $A$  and  $D$  can be extended to an eigenvector basis for  $M$ .

If  $M$  is diagonalizable, then the eigenvector basis for  $A$  and  $D$  can be used to construct an eigenvector basis for  $A$  and for  $D$ .

11. a

- [a]) Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be a  $2 \times 2$  matrix with some eigenvalue  $\lambda$ . Let

$$v = \begin{bmatrix} b \\ \lambda - a \end{bmatrix}$$

. We will show that  $v$  is an eigenvector for  $\lambda$  (assuming it is non-zero). This means that we should have

$$\begin{aligned} Av &= \lambda v && \iff \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} b \\ \lambda - a \end{bmatrix} &= \lambda \begin{bmatrix} b \\ \lambda - a \end{bmatrix} && \iff \\ \begin{bmatrix} ab - \lambda b - ab \\ bc + d\lambda - ad \end{bmatrix} &= \begin{bmatrix} \lambda b \\ \lambda(\lambda - a) \end{bmatrix} && \iff \\ \begin{bmatrix} \lambda b \\ bc + d\lambda - ad \end{bmatrix} &= \begin{bmatrix} \lambda b \\ \lambda(\lambda - a) \end{bmatrix} \end{aligned}$$

We see that the first coordinate is equal, whereas for the second, equality occurs if and only if

$$\begin{aligned} bc + d\lambda - ad &= \lambda(\lambda - a) \iff \\ \lambda^2 - (a + d)\lambda + (ad - bc) &= 0 \end{aligned}$$

Note that this is the characteristic polynomial evaluated at  $\lambda$ ! By definition,  $\lambda$  is a zero of the characteristic polynomial, and thus we conclude that the above inequality holds.

We conclude that  $v$  is an eigenvector for the value  $\lambda$ .

- [b]) Assume the matrix  $A$  has two distinct eigenvalues  $\lambda_1, \lambda_2$  and that  $b \neq 0$ . By [a], we have the two eigenvectors

$$\begin{bmatrix} b \\ \lambda_1 - a \end{bmatrix}, \begin{bmatrix} b \\ \lambda_2 - a \end{bmatrix}$$

Since  $b \neq 0$ , neither vector can be the zero vector. Since  $\lambda_1 \neq \lambda_2$ , we conclude that these are distinct eigenvectors of their respective eigenvalues.

A matrix  $P$  such that  $P^{-1}AP$  is given by the change of basis matrix with eigenvectors as columns:

$$\begin{bmatrix} b & b \\ \lambda_1 - a & \lambda_2 - a \end{bmatrix}$$

## Result

3 of 3

- Let  $A$  be the matrix and  $v$  the supposed eigenvector. By direct comparison, it is easy to check that  $Av = \lambda v$ .
- Assuming that there are two eigenvalues, part a) gives us an explicit way to write them down. The matrix  $P$  is given simply as a matrix with the two eigenvectors as columns.

## Section 7

- a

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

The characteristic polynomial is

$$p(t) = \begin{vmatrix} t-1 & -1 & 0 \\ 0 & t-1 & 0 \\ 0 & -1 & t-1 \end{vmatrix} = (t-1)^3$$

Thus the only eigenvalue is  $\lambda = 1$ . We search for the (generalized) eigenvectors i.e. vectors  $v$  from  $\ker(A - \lambda I)^j$  for  $j \geq 1$ . We find that

$$(A - \lambda I)v = 0 \iff \left[ \begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right]$$

We see that if  $v = (x_1, x_2, x_3)$ , we need to have  $x_2 = 0$ . Thus  $x_1, x_3$  can be chosen arbitrarily and we see that the  $\ker(A - \lambda I)$  is spanned e.g. by  $(1, 0, 0)$  and  $(0, 0, 1)$  and  $k_1$ , the dimension of this kernel, is equal to 2. Consider now  $(A - \lambda I)^2$ :

$$(A - \lambda I)^2 v = 0 \iff \left[ \begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Thus  $k_2 = \dim \ker(A - \lambda I) = 3$  and the same for higher powers.

Now,  $k_1$  tells us how many Jordan blocks there are of any size  $d \geq 1$ .  $k_2 - k_1$  tells us how many blocks of size  $d \geq 2$  there are and so on.

We see that there is exactly  $k_2 - k_1 = 1$  one block of size greater than 2. Since there are  $k_1 = 2$  blocks in total, we conclude that their sizes must be 1, 2. The eigenvalue  $\lambda$  is 1, so the Jordan normal form is given

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that the upper left part is a  $2 \times 2$  Jordan block. The 1 in lower right corner is a Jordan block of size 1.

## Result

2 of 2

We calculate the eigenvalues. It turns out that there is only one, namely  $\lambda = 1$ .

Consider  $(A - \lambda I)^j$ . We find that the dimensions of the kernel of this operator are  $k_1 = 2, k_2 = 3, \dots$ . This means that there are 2 Jordan blocks in total. There  $k_2 - k_1 = 1$  blocks of size greater than 2. The only possibility is that the two blocks are then of size 1 and 2.

2. a

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Then its square is

$$A^2 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} = A$$

which shows that  $A$  is idempotent. Suppose  $\lambda$  is an eigenvalue and suppose that  $v$  is a non-zero eigenvector. Then

$$\lambda v = Av = A^2v = \lambda^2 v$$

From this we have that  $(\lambda^2 - \lambda)v = 0$ . Since  $v$  has at least one non-zero coordinate, we conclude that  $\lambda$  must be either 0 or 1 (if there is an eigenvalue).

We try out  $\lambda = 0$ . The kernel of  $A - 0I = A$  is given by the solutions of

$$Ax = 0 \iff \left[ \begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ -1 & -1 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right]$$

Note that all three lines reduce to one and the same equation  $x_1 + x_2 + x_3 = 0$ . Thus taking e.g.  $(1, -1, 0)$  and  $(1, 0, -1)$  shows that the kernel of  $A$  is of dimension 2.

Now, let us check whether maybe 1 is an eigenvalue as well. We have

$$(A - I)x = 0 \iff \left[ \begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ -1 & -2 & -1 & 0 \\ 1 & 1 & 0 & 0 \end{array} \right]$$

This reduces to the two equations  $x_1 + x_2 = 0$  and  $x_2 + x_3 = 0$ . This space is spanned by e.g.  $(1, -1, 1)$ .

### Step 3

3 of 4

We see that we have found a basis given by eigenvectors, therefore the operator is diagonalizable and we have that its Jordan normal form consists of  $1 \times 1$  Jordan blocks (i.e. is a diagonal matrix)

$$J = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

### Result

4 of 4

It is easy to check that  $A^2 = A$ . From this it follows that an eigenvalue must satisfy  $\lambda^2 = \lambda$ . Thus the only possible eigenvalues, if there are any, are 0, 1. We see that for  $\lambda = 0$  we have two eigenvectors, while for  $\lambda = 1$  one. This gives a basis of eigenvectors, so  $A$  diagonalizable.

3. a

Let  $V$  be a complex vector space of dimension 5. Let  $T$  be a linear operator whose characteristic polynomial is  $(t - \lambda)^5$ . Suppose that the rank of  $T - \lambda I$  is equal to 2.

This means that

$$k_1 = \dim \ker(A - \lambda I) = 5 - 2 = 3$$

Recall that  $k_1$  states the **number** of Jordan blocks (of any size  $d \geq 2$ ). Since no other information is available, we conclude that Jordan normal form is made up of 3 blocks whose sizes can be partitioned as:

$$\begin{aligned} 5 &= 3 + 1 + 1 \\ &= 2 + 2 + 1 \end{aligned}$$

Thus the possible Jordan normal forms of the operator are:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(Note that the individual Jordan blocks could be permuted by a different choice of basis but there all similar matrices.)

## Result

2 of 2

Since the rank of  $T - \lambda I$  is equal to 2, the dimension of its kernel is  $k_1 = 3$ . This means that there are 3 Jordan blocks. Their sizes must add up to 5, thus the possibilities are 3, 1, 1 and 2, 2, 1.

4. a

- [a)] Consider matrices whose characteristic polynomial is given by

$$(t + 2)^2(t - 5)^3$$

Let  $M$  be such a matrix and suppose  $J$  is its normal form. It is immediately seen that  $-2$  can appear twice and 5 thrice (because of the multiplicities in the characteristic polynomial). Thus the possible Jordan blocks can be found by partitioning 2 and 3:

$$\begin{aligned} 2 &= 2 \\ &= 1 + 1 \\ 3 &= 3 \\ &= 2 + 1 \\ &= 1 + 1 + 1 \end{aligned}$$

The possible  $2 \cdot 3 = 6$  Jordan forms are given by

$$\begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}$$

$$\begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}$$



- [b] Assume now additionally that the space of eigenvectors with eigenvalue  $-2$  is one-dimensional, and the space of eigenvectors with eigenvalue  $5$  is two-dimensional.

This means that  $k_1 = \dim \ker(A + 2I) = 1$  and so there is only one Jordan block for  $\lambda = -2$ .

Similarly, for the other eigenvalue  $k_1 = \dim \ker(A - 5I) = 2$ . This means that there are two Jordan blocks for  $\lambda = 5$ .

This leaves the matrices:

$$\begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

## Result

3 of 3

a) From the characteristic polynomial we see that there are up to 2 Jordan blocks for  $\lambda = -2$  and 3 for  $\lambda = 5$ . In the first case, the possibilities are one block of size 2 or two of size 1. Similarly, for  $\lambda = 5$ , the sizes are 3, 2 + 1 and 1 + 1 + 1.

b) With the additional assumptions, only two matrices from a) satisfy the conditions.

5. a

Suppose a matrix  $A$  has the property that all the eigenvectors are multiples of a single vector  $v$ . Suppose  $w$  is another eigenvector. Then if  $v$  belongs to the eigenvalue  $\lambda$

$$Aw = A(kv) = kAv = k\lambda v = \lambda w$$

and so we see that  $w$  must also belong to the same eigenvalue  $\lambda$ . Thus there is only one eigenvalue. Since all the eigenvectors are multiples of  $v$ , we see that

$$k_1 = \dim \ker(A - \lambda I) = 1$$

This means that there is only one Jordan block i.e. the Jordan form is given by

$$\begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 0 \\ 0 & 0 & \dots & 1 & \lambda \end{bmatrix}$$

## Result

2 of 2

Since all the eigenvectors are multiples of one another, we conclude that all belong to one eigenvalue. Then  $k_1 = \dim \ker(A - \lambda I) = 1$  is the number of Jordan blocks. Thus the whole Jordan form is given by one block (a diagonal matrix with 1 on the left subdiagonal).

6. a



Suppose a linear operator  $T$  has a Jordan form  $J$  consisting of only one block. Suppose  $V$  was an invariant subspace. Consider its direct complement  $W$ . Then putting together bases for  $V$  and  $W$ , we would have that  $T$  has a representation as a block matrix

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

These matrices can be put into Jordan normal form (since  $T$  can), and we would find that the Jordan form  $J$  of  $T$  consists of at least two Jordan blocks. This is in contradiction with the starting hypothesis.

We conclude that the only invariant subspace  $T$  can have is the vector space on which it is defined.

## Result

2 of 2

Assume  $T$  has an invariant subspace smaller than the whole space on which it is defined. But then  $T$  can be represented as a (diagonal) block matrix. This is in contradiction with  $T$  having just one Jordan block.

7. a

Let  $A$  be a  $n \times n$  complex matrix such that  $A^2 = A$ . We show  $A$  is diagonalizable.

## Step 2

2 of 3

Let,  $\lambda \in \mathbb{C}$  be an eigenvalue of  $A$  and  $y \in \mathbb{C}^n$  is an eigenvector of  $A$  corresponding to the eigenvalue  $\lambda \in \mathbb{C}$  i.e.  $Ay = \lambda y$  i.e.  $\lambda y = Ay = A^2y = A(Ay) = A(\lambda y) = \lambda Ay = \lambda^2 y$ . i.e.  $(\lambda^2 - \lambda)y = 0$ . But  $y \neq 0$  as  $y$  is an eigenvector. Hence  $\lambda(\lambda - 1) = 0$ . So eigenvalues of  $A$  can be either 1 or 0.

Suppose, 0 is not an eigenvalue of  $A$  then,  $\det(A - 0I) \neq 0$  i.e.  $A$  is invertible. So the equation  $A^2 = A$  i.e.  $A(A - I) = 0$  implies  $A = I$ . Hence in this case  $A$  is diagonalizable.

Next let, 1 is not an eigenvalue of  $A$  then  $\det(A - 1I) \neq 0$  i.e.  $A - I$  is invertible. So the equation  $A^2 = A$  i.e.  $A(A - I) = 0$  implies  $A = 0$ . Hence in this case,  $A$  is diagonalizable.

Now suppose, both 0 and 1 are eigenvalues of  $A$ . Note that,  $A(Ax) = A^2x = Ax$  for each  $x \in \mathbb{C}^n$ . Also, for  $x \in \mathbb{C}^n$  we have  $x = Ax + (x - Ax)$ . Now  $A(x - Ax) = Ax - A(Ax) = 0, \forall x \in \mathbb{C}^n$ . Hence for each  $x \in \mathbb{C}^n$  we have  $Ax$  is either 0 or an eigenvector of  $A$  corresponding to eigenvalue 1. Similarly, for each  $x \in \mathbb{C}^n$  either  $x - Ax$  is zero or an eigenvector of  $A$  corresponding to eigenvalue 0. Hence,  $\mathbb{C}^n$  can be written as sum of two eigen-spaces of  $A$ . Now suppose,  $0 = u + v$  where,  $u \in \mathbb{C}^n$  is an eigenvector of  $A$  corresponding to eigenvalue 0 and  $v \in \mathbb{C}^n$  is an eigenvector of  $A$  corresponding to eigenvalue 1. Then  $0 = A0 = Au + Av = 0u + 1v = v$ . Hence  $u = 0$  also. Therefore,  $\mathbb{C}^n$  is direct sum of two eigen-spaces of  $A$ . Hence in this case  $A$  also diagonalizable.

## Result

We show that a  $n \times n$  matrix  $A$  with  $A^2 = A$  is diagonalizable.

8. a

Let  $A$  be a complex square matrix. Since it is over the complex numbers, the characteristic polynomial will have zeroes, and  $A$  will have eigenvalues. This means that  $A$  is similar to a matrix in Jordan normal form.

We will first prove the statement for matrices in Jordan normal form. In particular, consider first a single Jordan block:

$$J_\lambda = \begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix}$$

We claim that  $J_\lambda$  is similar to  $J_\lambda^t$ . Indeed, take as similarity matrix  $P$  the matrix with 1-s on the antidiagonal and 0-s else. Then

$$PJ = J^tP \iff \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix} = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix} \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Multiply the left and right hand side out. Note that the similarity matrix is a permutation matrix. Multiplying from the left it works on the rows, and from the right it works on the columns. Note the matrix just reverses the order of the rows/columns. We find that the two sides are equal:

$$\begin{bmatrix} 0 & 0 & \dots & 1 & \lambda \\ 0 & 0 & \dots & \lambda & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \lambda & \dots & 0 & 0 \\ \lambda & 0 & \dots & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 1 & \lambda \\ 0 & 0 & \dots & \lambda & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \lambda & \dots & 0 & 0 \\ \lambda & 0 & \dots & 0 & 0 \end{bmatrix}$$

We conclude that Jordan blocks  $J_\lambda$  are similar to upper triangular matrices.

Now, suppose  $J$  is a matrix in Jordan normal form. Then it consists of individual Jordan blocks

$$\begin{bmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_k \end{bmatrix}$$

For each of these  $J_i$  there is a similarity matrix  $P_i$  just as above. Let  $P$  the block matrix with the  $P_i$  in the diagonal. Then  $PJ = J^tP$  will again hold since it holds for all the Jordan blocks. Thus every matrix in Jordan normal form is similar to its transform.

Finally, since every matrix is similar to a matrix in Jordan normal form, we have a similarity matrix  $Q$ , such that

$$Q^{-1}AQ = J$$

Then multiplying with  $P^{-1}$  from the left and  $P$  from the right, one finds that

$$\begin{aligned} P^{-1}Q^{-1}AQ P &= P^{-1}JP \\ &= J^t \\ &= (QAQ^{-1})^t \\ &= (Q^{-1})^t A^t Q^t \end{aligned}$$

From which finally follows that for  $R = QP(Q^t)^{-1}$ , we have

$$R^{-1}AR = A^t$$

as required.

Thus every matrix over the complex numbers is similar to its transpose.

## Result

4 of 4

First prove that this is true for Jordan blocks  $J_\lambda$ . One can take for the similarity matrix the permutation matrix which reverses the order of rows/columns.

Next prove the statement for matrices in Jordan normal form. Finally, use the fact that every matrix is similar to a matrix in Jordan normal form to derive the general statement.

9. a

Here we have to find a  $2 \times 2$  matrix  $A$  over  $\mathbb{F}_p$  such that  $A^n = I$  for some positive integer  $n$  but  $A$  is not diagonalizable over  $\mathbb{F}_p$ .

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

. Then,

$$A^2 = \begin{bmatrix} 1 & 1+1 \\ 0 & 1 \end{bmatrix}$$

. Similarly,

$$A^3 = \begin{bmatrix} 1 & 1+1+1 \\ 0 & 1 \end{bmatrix}$$

. In general,

$$A^n = \begin{bmatrix} 1 & n \cdot 1 \\ 0 & 1 \end{bmatrix}$$

, where  $n \cdot 1 = 1 + 1 + 1 + \dots + 1 \{n\text{-times}\}$ . Therefore,

$$A^p = \begin{bmatrix} 1 & p \cdot 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Now we show  $A$  is not diagonalizable over  $\mathbb{F}_p$ . Note that, 1 is the only eigenvalue of  $A$ . So if possible let,  $A$  is diagonalizable, then eigenvectors of  $A$  corresponding to the eigenvalue 1 forms a basis of  $\mathbb{F}_p^2$ . But if

$$\begin{bmatrix} a \\ b \end{bmatrix}$$

is an eigenvector of  $A$  then

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 1 \begin{bmatrix} a \\ b \end{bmatrix}$$

i.e.

$$\begin{bmatrix} a+b \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

. So that,  $b = 0$ . So the eigen-space of  $A$  is

$$\left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} : x \in \mathbb{F}_p \right\}$$

which doesn't contain

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{F}_p^2$$

--- contradiction. Therefore,  $A$  can't be diagonalizable over  $\mathbb{F}_p$ .

## Result

Our required matrix is

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

## Miscellaneous Problem

1. a

Let  $v = (a_1, \dots, a_n)$  be a real row vector. We form the  $n! \times n$  matrix  $M$  whose rows are obtained by permuting the entries of  $v$  in all possible ways.

Note that if

$$a_1 = a_2 = \dots = a_n$$

then the rank of the matrix  $M$  is 0 (if all the entries are zero) or 1 otherwise. We will see that in general this case is quite important.

Assume now that not all the  $a_i$  are equal. Let  $V \subseteq \mathbb{R}^n$  be the vector space spanned by the row vectors of  $M$ . Define

$$V^\perp = \{x \in \mathbb{R}^n \mid Mx = 0\} = \ker M$$

(this is in fact the orthogonal complement of  $V$ ). We claim that one of these two subspaces must be equal to

$$\text{span}\{(1, 1, \dots, 1)\} = \{(c, c, \dots, c) \mid c \in \mathbb{R}\}$$

Suppose this was not the case. Then both  $V$  and  $V^\perp$  have vectors whose components are not all the same. Let

$$w = (c_1, c_2, \dots, c_n) \in V^\perp$$

with  $c_i \neq c_j$ . We assumed that not all the  $a_i$  are the same. Since  $V$  contains all permutations, without loss of generality we may assume that  $a_i \neq a_j$  (else reshuffle the vector  $v$ ). Define the vector

$$v' = (a_1, \dots, a_j, \dots, a_i, \dots, a_n)$$

which has the  $i$ -th and  $j$ -th components of  $v$  switched. Since  $M$  contains all the permutations, this is still a row vector of  $M$ . Finally note that because of  $Mw = 0$ , we have  $w^t v = w^t v' = 0$  so that

$$0 = w^t v - w^t v' = w^t (v - v')$$

$v - v'$  has only two non-zero components, those in positions  $i$  and  $j$ . In fact:

$$w^t (v - v') = (c_i - c_j)(a_i - a_j)$$

By assumption  $c_i \neq c_j$  and  $a_i \neq a_j$ , so we have finally arrived at a contradiction. Since we have assumed that not all of the  $a_i$  are the same, we conclude that

$$V^t = \{(c, c, \dots, c) \mid c \in \mathbb{R}\}$$

Now,  $V$  is spanned by the column vectors - this means that  $V = M(\mathbb{R}^n)$ . On the other hand  $V^t = \ker M$ . The dimensions of these two spaces must sum to  $n$ . Since  $V^t$  is of dimension 1 or 0, we see that the dimension of  $V$  is either  $n - 1$  or  $n$ .

Thus the possible ranks are 0, 1,  $n - 1$ ,  $n$ . These are achieved by taking e.g.

- for rank 0 take  $v = (0, 0, \dots, 0)$
- for rank 1 take  $v = (1, 1, \dots, 1)$
- for rank  $n - 1$  take  $v = (1, -1, \dots, 0)$
- for rank  $n$  take  $v = (1, 0, \dots, 0)$

## Result

4 of 4

Let  $V$  be the vector space spanned by the columns. This is in fact the image of the matrix  $M$  on  $\mathbb{R}^n$ . We consider the kernel as well and show that one of the two must be the space spanned by  $(1, 1, \dots, 1)$ . This shows that either

$$a_1 = a_2 = \dots = a_n \text{ or } a_1 + a_2 + \dots + a_n = 0. \text{ This gives the possible ranks } 0, 1, n - 1, n.$$

2. a



- [a] Let  $A$  be a complex  $n \times n$  matrix with distinct eigenvalues  $\lambda_1, \dots, \lambda_n$ . Suppose  $|\lambda_1| > |\lambda_i|$  for  $i > 1$ .

Let  $X$  be some vector in  $\mathbb{C}^n$ . Suppose that  $v_1, \dots, v_n$  are eigenvectors of the corresponding eigenvalues. Since there are  $n$  of them, they form a basis for  $\mathbb{C}^n$  and so  $X$  can be written as a linear combination of them. Using this, we find that:

$$\begin{aligned} X_k &= \lambda_1^{-k} A^k X \\ X_k &= \lambda_1^{-k} A^k (a_1 v_1 + \dots + a_n v_n) \\ X_k &= \lambda_1^{-k} (a_1 \lambda_1^k v_1 + \dots + a_n \lambda_n^k v_n) \\ X_k &= a_1 \left( \frac{\lambda_1}{\lambda_1} \right)^k v_1 + \dots + \left( \frac{\lambda_n}{\lambda_1} \right)^k v_n \quad / \quad \lim_{k \rightarrow \infty} \\ X_k &= a_1 v_1 \end{aligned}$$

since  $(\lambda_i/\lambda_1)^k$  tends to zero for all the  $i > 1$ .

Note that the process doesn't work if and only if  $v$  is from the  $n - 1$  dimensional space spanned by the other eigenvalues  $v_2, \dots, v_n$ .

Let  $v = a_1 \lambda_1^k v_1 + \dots + a_n \lambda_n v_n$  be a randomly chosen vector. The probability that  $a_1 = 0$  is zero. Thus the method should work for most randomly chosen vectors.

- [b] Let  $A$  be a complex  $n \times n$  matrix with non-necessarily distinct eigenvalues  $\lambda_1, \dots, \lambda_n$ . Suppose  $|\lambda_1| > |\lambda_i|$  for  $i > 1$  (so the largest eigenvalue still is unique).

Let  $X$  be some vector in  $\mathbb{C}^n$ . Suppose that  $v_1, \dots, v_n$  are the generalized eigenvectors of the corresponding eigenvalues. Since there are  $n$  of them, they form a basis for  $\mathbb{C}^n$  and so  $X$  can be written as a linear combination of them. We also invoke the Jordan normal form. Let  $P$  be the similarity matrix which sends the standard basis to the (generalized) eigenvector basis. Then

$$\begin{aligned} J &= P^{-1} A P \\ A &= P J P^{-1} \end{aligned}$$

Note that  $P^{-1}$  sends the (generalized) eigenvector basis  $(v_i)$  to the standard basis  $(e_i)$ . Thus:

$$\begin{aligned} X_k &= \lambda_1^{-k} A^k X \\ X_k &= \lambda_1^{-k} (P J P^{-1})^k (a_1 v_1 + \dots + a_n v_n) \\ X_k &= \lambda_1^{-k} P J^k P^{-1} ((a_1 v_1 + \dots + a_n v_n)) \\ X_k &= \lambda_1^{-k} P J^k (a_1 e_1 + \dots + a_n e_n) \end{aligned}$$



- [b)]

Now consider the matrix  $\lambda_1^{-k} J^k$ . It is of the form:

$$\begin{bmatrix} 1 & & & \\ & \lambda_1^{-k} J_{\lambda_2}^k & & \\ & & \ddots & \\ & & & \lambda_1^{-k} J_{\lambda_n}^k \end{bmatrix}$$

Decompose the  $J_\lambda$  into the diagonal and nilpotent part  $J_\lambda = \lambda I + N$ . Then

$$\lambda_1^{-k} J_\lambda^k = \left( \frac{\lambda_2}{\lambda_1} \right)^k \sum_{i=0}^k \binom{k}{i} (\lambda_2^{-1} N)^i$$

If  $k$  is large enough ( $k > n$ ), then the right sum will still have only  $n$  terms since  $N^n = 0$ . Taking the limit, we find that

$$\begin{aligned} \lim_{k \rightarrow \infty} \lambda_1^{-k} J_\lambda^k &= \lim_{k \rightarrow \infty} \left( \frac{\lambda_2}{\lambda_1} \right)^k \sum_{i=0}^n \binom{k}{i} (\lambda_2^{-1} N)^i \\ &= \sum_{i=0}^n \lim_{k \rightarrow \infty} \left( \frac{\lambda_2}{\lambda_1} \right)^k \binom{k}{i} (\lambda_2^{-1} N)^i \\ &= 0 \end{aligned}$$

Note that the binomial coefficients are polynomials in  $k$  and therefore grow slower compared to the exponential  $\left( \frac{\lambda_2}{\lambda_1} \right)^k$  tending to zero. Therefore

$$\lim_{k \rightarrow \infty} J^k = \lim_{k \rightarrow \infty} \begin{bmatrix} 1 & & & \\ & \lambda_1^{-k} J_{\lambda_2}^k & & \\ & & \ddots & \\ & & & \lambda_1^{-k} J_{\lambda_n}^k \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}$$

- [b)]

Now go back to  $X_k$ .

$$\lim_{k \rightarrow \infty} X_k = P \left( \lim_{k \rightarrow \infty} \lambda_1^{-k} J^k \right) (a_1 e_1 + \dots + a_n e_n)$$

As shown above the limit will tend to a matrix which has only the entry 1 in top left corner. Thus

$$\begin{aligned} \lim_{k \rightarrow \infty} X_k &= P a_1 e_1 \\ &= a_1 v_1 \end{aligned}$$

giving the eigenvector for the eigenvalue  $\lambda_1$ .

## Result

5 of 5

a) Take a basis of eigenvectors, and write  $X$  in that basis. Note that  $(\lambda_i/\lambda_1)^k$  tends to zero.

b) Take a basis of generalized eigenvectors. Write  $X$  in that basis. Let  $P$  the similarity matrix with the Jordan normal form of  $A$ . Prove that  $(\lambda_1^{-1} J)^k$  tends to a matrix with only 1 entry in its top left corner. Then one will see that  $X_k$  tends to  $v_1$  as it should.

3. a

In the previous exercise, we proved that if  $\lambda_1$  is an eigenvalue such that  $|\lambda_1| > |\lambda_i|$  for all  $i > 1$ , then

$$X_k = \lambda_1^{-k} A^k X$$

converges to an eigenvector  $v$  for  $\lambda_1$ . This is great if we know  $\lambda_1$ , but here we are asked to calculate it by such a method.

Our matrix is

$$A = \begin{bmatrix} 3 & 1 \\ 3 & 4 \end{bmatrix}$$

. We would like to approximate the size of  $\lambda_1$  in some way without calculating it explicitly. Note that since  $X_k$  is close to an eigenvector, we may naively expect  $\lambda_1^k X_k$  to be close to an eigenvalue as well. Thus we calculate the sequence  $A^k X$  for some vector  $X$ . Take  $X$  to be something simple e.g.  $X = (1, 0)$ . Then we have

$$\begin{aligned} \lambda_1 X_1 &= AX = (3, 3) \\ \lambda_1^2 X_2 &= A^2 X = (12, 21) \\ \lambda_1^3 X_3 &= A^3 X = (57, 120) \\ \lambda_1^4 X_4 &= A^4 X = (231, 651) \\ \lambda_1^5 X_5 &= A^5 X = (1524, 3477) \end{aligned}$$

Recall that any multiple of an eigenvector is an eigenvector.

We try to scale these vectors - the easiest way is to divide out the second component.

$$\begin{aligned} \lambda_1 X_1 &= AX = 3(1, 1) \\ \lambda_1^2 X_2 &= A^2 X = 21(0.5714, 1) \\ \lambda_1^3 X_3 &= A^3 X = 120(0.475, 1) \\ \lambda_1^4 X_4 &= A^4 X = 651(0.4470, 1) \\ \lambda_1^5 X_5 &= A^5 X = 3477(0.4383, 1) \\ \lambda_1^6 X_6 &= A^6 X = 18480(0.4355, 1) \\ \lambda_1^7 X_7 &= A^7 X = 98067(0.4347, 1) \\ \lambda_1^8 X_8 &= A^8 X = 520149(0.4344, 1) \end{aligned}$$

Further iterations yield the same 3 decimal places accuracy. Thus  $v = (0.4344, 1)$  should be extremely close to an eigenvector with the largest eigenvalue. We check:

$$\begin{aligned} Av &= \lambda_1 v \\ A(2.3032, 5.3032) &= \lambda_1(0.4344, 1) \end{aligned}$$

Thus  $\lambda_1$  should be approximately 5.3032 and  $\frac{2.3032}{0.4344} \approx 5.302$ . Averaging the two, we get accuracy to three decimals, and find that  $\lambda_1 = 5.303$ .

## Result

3 of 3

Calculate  $A^j X$  for some random vector  $X$ . This should be approximately  $\lambda_1^j v$  where  $v$  is an eigenvector. Scaling the vectors, we should find that they converge to an eigenvector. Calculate the eigenvector to sufficient precision, and then find the eigenvalue from that.

4. a

Let  $A$  be an infinite real matrix. Consider the matrix product  $XA$  where  $X$  is an infinite row vector. If this product is to be defined for **all** row vectors, then we would need that

$$\begin{bmatrix} x_1 & x_2 & x_3 & \dots \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots \\ \vdots \end{bmatrix}$$

Thus we would need all the series

$$\sum_{k=1}^{\infty} a_{ik}x_k$$

to converge for each row  $i$  (and every vector  $x$ ). Fix an  $i$ . We choose a vector which will surely make the sum infinite if there are an infinite of non-zero  $a_{ik}$ . This can be achieved by taking:

$$x_k = \frac{1 + a_{ik}^2}{a_{ik}}$$

whenever the expression is defined (if  $a_{ik} = 0$  defined  $x_k$  can be whatever e.g. 0). Then the  $i$ -th row of  $XA$  will have

$$\begin{aligned} \sum_{k=1}^{\infty} a_{ik}x_k &= \sum_{k=1}^{\infty} a_{ik}x_k \\ &= \sum_{\substack{k=1 \\ a_{ik} \neq 0}}^{\infty} a_{ik} \frac{a_{ik}^2 + 1}{a_{ik}} \\ &= \sum_{\substack{k=1 \\ a_{ik} \neq 0}}^{\infty} (a_{ik}^2 + 1) \end{aligned}$$

If there are infinitely many non-zero  $a_{ik}$ , then the sum will be infinite. Since  $i$  was arbitrary, this holds for every row  $i$ .

We conclude that if  $XA$  is to be well-defined on the whole space of infinite row vectors, we must have that every column of  $A$  must have only finitely many non-zero entries.

Now consider the same thing but on the space

$$Z = \{x \mid x \text{ has finitely many non-zero entries}\}$$

Then the sum

$$\sum_{k=1}^{\infty} a_{ik}x_k$$

always reduces to a finite sum which is always well-defined to evaluate. Thus  $XA$  is defined for any matrix  $A$  on the space  $Z$ .

## Result

3 of 3

Note that  $XA$  gives rise to a vector whose every is a series. If  $X$  can be arbitrary, then it's not hard to choose an  $X$  such that at least one of the series diverges. The only possibility is if every column has only finitely many non-zero entries.

On the other hand, on the space  $Z$ , each entry in  $XA$  is a finite sum and so the product is always well-defined.

## 5. a

Let  $F$  be a field and  $A$  be a  $m \times n$  matrix over  $F$ . Let  $\phi : F^n \rightarrow F^m$  is the linear operator defined by  $\phi(x) = Ax$ ,

$\forall x \in F^n$ . We have to show the following are equivalent :-

1. There is a  $n \times m$  matrix  $B$  with  $AB = I_{m \times m}$ .
2.  $\phi$  is surjective.
3. The rank of  $A$  is  $m$ .

1  $\implies$  2.

So let  $y \in F^m$ . We have to show there is  $x \in F^n$  such that  $\phi(x) = y$ . Let  $A_{s1}, A_{s2}, \dots, A_{sn}$  denote the columns of  $A$ .

Also let  $e_{s1}, e_{s2}, \dots, e_{sm}$  denote the columns of  $I_{m \times m}$ . Now  $AB = I_{m \times m}$  gives that,  $\sum_{j=1}^n b_{jk} A_{sj} = e_{sk}, \forall k = 1, 2, \dots, m$ .

Here,  $b_{jk}$  denotes the entry of  $B$  in the intersection of  $j$ -th row and  $k$ -th column of  $B$ . This shows that,

each column of  $I_{m \times m}$  is in the column space of  $A$ . But note that  $F^m = \text{span}\{e_{s1}, \dots, e_{sm}\}$ .

Hence column space of  $A$  is nothing but  $F^m$ . So that,  $y \in F^m$  can be written as  $y = \sum_{j=1}^n \lambda_j A_{sj}$ . Hence  $\phi(\lambda_1, \dots, \lambda_n)^t = y$ .

2.  $\implies$  3.

$\phi$  is surjective implies for each  $y \in F^m$  there is  $x \in F^n$  such that  $Ax = \phi(x) = y$  i.e.

Hence column space of  $A$  is  $F^m$ . So the, rank of  $A$  = column rank of  $A = m$ .

3.  $\implies$  1.

Suppose the rank of  $A$  is  $m$ . But rank  $A$  is same dimension of column space of  $A$ .

Therefore, dimension of column space is  $m$ . Note that each column of  $A$  is an element of  $F^m$  and  $\dim(F^m) = m$ .

So column space of  $A$  is same as  $F^m$ . In particular,  $e_{sk}$  can be written as  $\sum_{j=1}^n b_{jk} A_{sj} = e_{sk}$  for some scalars  $b_{jk}$  for all  $j = 1, 2, \dots, n$  and  $k = 1, 2, \dots, m$ .

Here,  $A_{sj}$  denotes the  $j$ -th column of  $A$  and  $e_{sk}$  denotes  $k$ -th column of  $I_{m \times m}$ .

Consider the  $n \times m$  matrix  $B$  whose  $(j, k)$ -th entry is  $b_{jk}$ , for all  $j = 1, \dots, n$  and  $k = 1, \dots, m$ .

Let  $F$  be a field and  $A$  be a  $m \times n$  matrix over  $F$ . Let  $\phi : F^n \rightarrow F^m$  is the linear operator defined by  $\phi(x) = Ax, \forall x \in F^n$ .

We have to show the following are equivalent :-

1. There is a  $n \times m$  matrix  $B$  with  $BA = I_{n \times n}$ .
  2.  $\phi$  is injective.
  3. The rank of  $A$  is  $n$ .
1.  $\implies$  2.

Suppose,  $\phi(x) = 0$  for some  $x \in F^n$ . Therefore,  $Ax = 0$ . Now  $BA = I_{n \times n}$  implies that  $0 = B0 = B(Ax) = I_{n \times n}x = x$ . So that  $\phi$  is injective.



2.  $\implies$  3.

$\phi$  is injective implies null space of  $\phi$  is  $\{0\}$ . Therefore dimension formula says,  $0 + \text{rank}(\phi) = \text{nulity}(\phi) + \text{rank}(\phi) = \dim(F^n) = n$ . That's  $\text{rank}(\phi) = n$ .

Hence range space of  $\phi$  is  $F^n$ . Therefore, column space of  $A$  is  $F^n$ . Hence column rank of  $A$  is  $n$ .

But column rank is same as rank. Therefore, rank of  $A$  is  $n$ .

3.  $\implies$  1.

Suppose, rank of  $A$  is  $n$ . But row rank is same as column rank. Hence dimension of row space of  $A$  is  $n$ .

Note that,  $i$ -th row  $A_{i*}$  of  $A$  is an element of  $F^n$ , for all  $i = 1, \dots, m$ . But  $\dim(F^n) = n$ .

Hence row space of  $A$  is same as  $F^n$ . In particular,  $i$ -th row  $e_{i*}$  of  $I_{n \times n}$  can be written as  $e_{i*} = \sum_{t=1}^m b_{it} A_{t*}$  for each  $i = 1, \dots, n$ .

Let  $B$  be the  $n \times n$  matrix whose  $(i, t)$ -entry is  $b_{it}$  for each  $i = 1, \dots, n$  and for each  $t = 1, \dots, m$ . Hence,  $BA = I_{n \times n}$ .

## Result

We use the fact that row rank of a matrix is same as column rank.

6. a

We have to show a linear operator  $T$  on a vector space  $V$  over the field  $F$  of dimension  $n$

can have at most  $n$  distinct eigenvalues **without using characteristic polynomial**.

So let  $\lambda_1, \dots, \lambda_m$  be a set of distinct eigenvalues of  $T$ . Also let,  $v_1, \dots, v_m$  be a collection of eigenvectors of  $T$  such that  $T(v_k) = \lambda_k v_k, \forall k = 1, \dots, m$ .

Suppose,  $v_i = v_j$  for some  $i, j \in \{1, \dots, m\}$ . Then  $T(v_i) = T(v_j)$  i.e.  $\lambda_i v_i = \lambda_j v_j$  i.e.  $(\lambda_i - \lambda_j)v_i = 0$ . Now  $v_i$  is an eigenvector, so  $v_i \neq 0$  i.e.  $\lambda_i = \lambda_j$ . Hence, the collection  $v_1, \dots, v_m$  is actually a collection of distinct eigenvectors.

Now for a polynomial  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_lx^l \in F[x]$  we define  $f(T) := a_0I + a_1T + a_2T^2 + \dots + a_lT^l$ . So that, for an eigenvalue  $\lambda$  of  $T$  with eigenvector  $v$  we have,  $f(T)(v) = a_0I(v) + a_1T(v) + a_2T^2(v) + \dots + a_lT^l(v) = (a_0 + a_1\lambda + a_2\lambda^2 + \dots + a_l\lambda^l)v = f(\lambda)v$ .

For each  $k = 1, \dots, m$  we define a polynomial  $f_k(x) \in F[x]$  such that  $f_k(\lambda_j) = 0$  if  $k \neq j$  and  $f_k(\lambda_j) = 1$  if  $k = j$ . So define,

$$\begin{aligned} f_k(x) &= \frac{(x - \lambda_1) \dots (x - \lambda_{k-1})(x - \lambda_{k+1}) \dots (x - \lambda_m)}{(\lambda_k - \lambda_1) \dots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \dots (\lambda_k - \lambda_m)} \\ &= \frac{\prod_{i=1, i \neq k}^m (x - \lambda_i)}{\prod_{i=1, i \neq k}^m (\lambda_k - \lambda_i)}. \end{aligned}$$

Hence  $f_k(T)(v_j) = f(\lambda_j)v_j = 0$  if  $j \neq k$  and  $f_k(T)(v_j) = f(\lambda_j)v_j = v_j$  if  $k = j$ .

Now suppose, for some scalars  $c_1, \dots, c_m$  we have  $c_1v_1 + \dots + c_mv_m = 0$ . So that,  $0 = f_k(T)(0) = f_k(T)(c_1v_1 + \dots + c_mv_m) = c_1f_k(T)v_1 + \dots + c_mf_k(T)v_m = (c_k \cdot 1)v_k = c_kv_k$ . Now  $v_k \neq 0$  as  $v_k$  is an eigenvector. So that,  $c_k = 0$ , for each  $k = 1, \dots, m$ . Therefore,  $\{v_1, \dots, v_m\}$  is a linearly independent set. But dimension of  $V$  is  $n$ . So that,  $m \leq n$ . Hence,  $T$  can have at most  $n$ -many distinct eigenvalues.

Using Lagrange interpolation polynomials we have shown that, any set of  $m$  distinct eigenvectors corresponding to distinct  $m$  eigenvalues is linearly independent.

## 7. a

(a)

To show that  $K_1 \subset K_2 \subset \dots$  and  $W_1 \supset W_2 \supset \dots$

Suppose  $V$  be a  $n$ -dimensional vector space,  $T$  be the linear operator on  $V$  and  $K_r$  and  $W_r$  be the kernel and images respectively of  $T^r$ .

Since, there arise two cases.

Case-1, when the matrix of the linear operator is nonsingular and upper triangular

Then,

$$|T| \neq 0$$

This implies that  $T^{-1}$  exists.

Then, the situation is trivial

Case-2, when the matrix of the linear operator is singular and strict triangular.

Then, dimension of  $T^r$  will decrease as  $r$  will increase and will vanish at  $r = n$ .

In this case, the number of elements in kernel  $K_r$  will increase for  $T^r$  as  $r$  will increase and number of elements in image  $W_r$  will decrease for  $T^r$  as  $r$  will increase.

Then, for any  $x \in K_1$

Then,

$$T(x) = 0$$

And

$$x \in K_2$$

This implies that,

$$K_1 \subseteq K_2$$

But, for any  $y \in K_2$ ,  $T(y) \neq 0$  in  $K_1$

Then,

$$y \notin K_1$$

And so on in the similar manner,

$$K_1 \subset K_2 \subset \dots$$

Since, both  $K_r$  and  $W_r$  are the subspaces of the vector space  $V$ .

And

$$\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$$

Then,

$$\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$$

$$\dim(\operatorname{im} T) = n - \dim(\ker T)$$

Since,  $T$  is linear operator on the vector space  $V$ , then  $T^r$  will also be the linear operator on the vector space  $V$

Then,

$$\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$$

$$\dim(\operatorname{im} T) = n - \dim(\ker T)$$

$$\dim(W_r) = n - \dim(K_r)$$

Then,  $r$  will increases the number of elements in kernel  $K_r$  will increase.

Then, the number of elements in image  $W_r$  will decrease.



Then,

$$W_1 \supset W_2 \supset \dots$$

Hence, **it concludes that kernels  $K_1 \subset K_2 \subset \dots$  and images  $W_1 \supset W_2 \supset \dots$ .**

(b)

To find all implications among the conditions (1)-(4),

Suppose  $V$  be a  $n$ -dimensional vector space,  $T$  be the linear operator on  $V$  and  $K_r$  and  $W_r$  be the kernel and images respectively of  $T^r$ .

Then, the implications on the condition (1),

$$K_r = K_{r+1}$$

Then, the restriction at  $T$  is that the linear operator  $T$  is restricted to the image  $W_r$  is an automorphism.

This implies that,

$$W_r = W_{r+1}$$

Hence, the required implication on the condition (1) is that  $\boxed{W_r = W_{r+1}}$ .

To find the implication on the condition (2),

Then, the implications on the condition (2),

$$W_r = W_{r+1}$$

This implies that, the restriction at  $T$  is that the linear operator  $T$  is restricted to the image  $W_r$  is an automorphism.

Hence, **the required implication on the condition (2) is that  $T$  is restricted to the image  $W_r$  is an automorphism.**

To find the implication on the condition (3),

Then, the implications on the condition (3),

$$W_r \cap K_1 = \{0\}$$

Then, for any  $v \in W_r \cap K_1$

This implies that,

$$T(v) = 0$$

This implies that,

$$v = 0$$

Then, the restriction at  $T$  is that the linear operator  $T$  is restricted to the image  $W_r$  is an automorphism.

Then,

$$W_r = W_{r+1}$$

Hence, the required implication on the condition (1) is that  $\boxed{W_r = W_{r+1}}$ .

To find the implication on the condition (4),

Then, the implications on the condition (4),

$$W_i + K_r = V$$

Since,

$$T(K_{r+1}) \subseteq K_r$$

Then, there is a induced map  $T' : \frac{V}{K_{r+1}} \rightarrow \frac{V}{K_r}$

Then,

$$T'(v') = 0$$

This implies that if and only if,

$$T(v) \in K_r$$

This implies that if and only if,

$$v' = 0$$

This implies that  $T'$  is injective map.

And for any  $v \in V$  and for some  $v' \in V$ , and for some  $v'' \in K_r$

$$v = T(v') + v''$$

Then,  $T'$  is surjective if and only if for any  $v \in V$ ,

$$v = T(v') + v''$$

This implies that if and only if,

$$W_i + K_r = V$$

Since,

$$\dim\left(\frac{V}{K_{r+1}}\right) = \dim\left(\frac{V}{K_r}\right)$$

This implies that,

$$\dim(K_{r+1}) = \dim(K_r)$$

This implies that,

$$K_{r+1} = K_r$$

This implies that,  $T'$  is surjective

Then, the restriction at  $T$  is that the linear operator  $T$  is restricted to the image  $W_r$  is an automorphism.

This implies that,

$$W_r = W_{r+1}$$

Hence, the required implication on the condition (4) is that  $\boxed{W_r = W_{r+1}}$ .

(c)

To find all implications among the conditions (1)-(4) for infinite dimensional vector space  $V$ ,

Suppose  $V$  be a infinite dimensional vector space,  $T$  be the linear operator on  $V$  and  $K_r$  and  $W_r$  be the kernel and images respectively of  $T^r$ .

Then, the implications on the condition (1),

$$K_r = K_{r+1}$$

This implies that,

$$K_r = K_{r+1} = \{0\}$$

And the linear operator  $T : W_r \rightarrow W_{r+1}$

Since,

$$K_r = K_{r+1} = \{0\}$$

Then, the linear operator  $T$  is injective.

Hence, **the required implication on the condition (1) is that the linear operator  $T$  on the infinite dimensional vector space  $V$  is injective.**

Since, whenever

$$K_r = K_{r+1} = \{0\}$$

This does not implies that,

$$W_r = W_{r+1}$$

Then, there is no linear operator that makes the implication between the condition (1) and (2).

Hence, **there is no implication between the condition (1) and (2).**

To find the implication on the condition (3),

Then, the implications on the condition (3),

$$W_r \cap K_1 = \{0\}$$

Then, for any  $v \in W_r \cap K_1$

This implies that,

$$T(v) = 0$$

This implies that,

$$v = 0$$

Then, the restriction at  $T$  is that the linear operator  $T$  is restricted to the image  $W_r$  is an automorphism.

Then,

$$W_r = W_{r+1}$$

Then, the linear operator  $T$  is injective.

Then,

$$K_r = K_{r+1} = \{0\}$$

Hence, **it concludes that the implication on the condition (1) and the implication on the condition (3) are equivalent.**

To find the implication on the condition (4),

Then, the implications on the condition (4),

$$W_1 + K_r = V$$

Since,

$$T(K_{r+1}) \subseteq K_r$$

Then, there is a induced map  $T' : \frac{V}{K_{r+1}} \rightarrow \frac{V}{K_r}$

Then,

$$T'(v') = 0$$

This implies that if and only if,

$$T(v) \in K_r$$

This implies that if and only if,

$$v' = 0$$

This implies that  $T'$  is injective map.

And for any  $v \in V$  and for some  $v' \in V$ , and for some  $v'' \in K_r$

$$v = T(v') + v''$$

Then,  $T'$  is surjective if and only if for any  $v \in V$ ,

$$v = T(v') + v''$$

This implies that if and only if,

$$W_1 + K_r = V$$

Since,

$$\dim\left(\frac{V}{K_{r+1}}\right) = \dim\left(\frac{V}{K_r}\right)$$

This implies that,

$$\dim(K_{r+1}) = \dim(K_r)$$

This implies that,  $T'$  is surjective map.

Since,  $T'$  is also injective map.

Therefore, the induced map  $T'$  is isomorphism from  $\frac{V}{K_{r+1}} \rightarrow \frac{V}{K_r}$ .

Since,  $K_r$  and  $W_r$  be the kernel and images for the map  $T' : V \rightarrow V$ .

Then, there is an induced map  $T'^r$ , which is an isomorphism as,

$$T'^r : \frac{V}{K_r} \rightarrow W_r$$

In the same way, the induced map  $T'^{(r+1)r}$

$$T'^{(r+1)r} : \frac{V}{K_{r+1}} \rightarrow W_{r+1}$$

Then, the induced map  $T'^{(r+1)r}$  is an isomorphism.

Then, the image space  $W_{r+1}$  is isomorphic to  $\frac{V}{K_{r+1}}$ , the space  $\frac{V}{K_{r+1}}$  is isomorphic to  $\frac{V}{K_r}$ , this

$\frac{V}{K_r}$  is isomorphic to the space  $W_r$ .

This implies that, the image space  $W_{r+1}$  is isomorphic to the image space  $W_r$ .

Then, the conclusion is that  $W_r = W_{r+1}$  if and only if  $T' : \frac{V}{K_{r+1}} \rightarrow \frac{V}{K_r}$  is surjective.

Hence, it concludes that the implication on the condition (2) and the implication on the condition (4) are equivalent.

## Method 2.

Let  $T$  be a linear operator on a vector space  $V$  and for each positive integer let  $K_r := \ker$

Let  $r$  be a positive integer and  $v \in K_r$  i.e.  $T^r(v) = 0$ . Then  $T^{r+1}(v) = T(T^r(v)) = T(0) = 0$ . So that,  $v \in K_{r+1}$ . Since  $v$  is

Now let  $s$  be a positive integer and  $T^{s+1}(v)$  be an arbitrary element of  $W_{s+1} = \text{im}(T^{s+1})$ . Then  $T^{s+1}(v)$

∗). Since  $v$  is arbitrary in  $K_r$  we have  $K_r \subseteq K_{r+1}$ . Now  $r$  is also arbitrary positive integer. So we have  $K_1 \subseteq K_2 \subseteq K_3 \subseteq \dots$  then  $T^{s+1}(v) = T^s(T(v)) \in \text{im}(T^s) = W_s$ . Since  $s$  is arbitrary positive integer we have,  $W_1 \supseteq W_2 \supseteq W_3 \supseteq \dots$

Now let  $T$  be a linear operator

Now assume (I) and let  $v \in W_r \cap K_1$ . Then  $v = T^r(u)$  for some  $u$

Now suppose (4). And choose  $T^r(v) \in W_r$ . Since  $v \in$

operator on a **finite dimensional** vector space  $V$ . Suppose  $r$  is a fixed positive integer. We show that the following are equivalent:

positive integer  $n$ . Now  $K_r = K_{r+1}$  if and only if  $\dim(K_r) = \dim(K_{r+1})$  if and only if  $\dim(V) - \dim(K_r) = \dim(V) - \dim(K_{r+1})$  if and only if for some  $u \in V$  and  $T(v) = 0$ . Therefore,  $0 = T(v) = T(T^{r-1}(u)) = T^{r-1}(u)$  i.e.  $u \in K_{r-1}$ . By assumption, we have  $u \in K_r$  also. Hence,  $u \in K_r \cap K_{r-1}$ . Since here are finite dimensional we have,  $\dim(\text{im}(T|_{W_r})) + \dim(\ker(T|_{W_r})) = \dim(W_r)$ . But  $W_r \supseteq W_{r+1}$  implies  $T|_{W_r}$  is surjective. Since  $v \in V = W_1 + K_r$ , we have  $v = T(x) + y$  for  $T(x) \in W_1$  and  $y \in K_r$ . Then  $T^r(v) = T^r(T(x)) + T^r(y) = T^{r+1}(x) + 0 = T^{r+1}(x)$  is a linear operator. But using the assumption we have,  $\dim(G) = \dim(V) - \dim(K_r) = \dim(V) - \dim(K_{r+1}) = \dim(H)$ . So that,  $|T|_H$

if and only if  $\dim(W_r) = \dim(W_{r+1})$  if and only if  $W_r = W_{r+1}$ . So (1) and (2) are equivalent.

surjective. So  $\dim(W_{r+1}) = \dim(\text{im}(T|_{W_r})) = \dim(W_r)$ . Hence  $W_r = W_{r+1}$ . This shows (3) implies (2).

$H$  is invertible operator. Now let,  $v \in V = G \oplus K_r$  with  $v = g + u$  for  $g \in G$  and  $u \in K_r$ . Since  $T|_H$  is surjective we have,  $h \in$

$h \in H$  such that  $T(h) = T|_H(h) = g$ . Hence  $v = T(h) + u$ . That's  $V = W_1 + K_T$ . This shows, (1) implies (4).

64







Let  $T$  be a linear operator on a finite-dimensional complex vector space  $V$ .

- [a] Let  $\lambda$  be an eigenvalue of  $V$  and let  $V_\lambda$  be the set of generalized eigenvectors (with the zero vector).

The generalized eigenvectors belonging to  $\lambda$  are given by those vectors  $v$  satisfying  $(T - \lambda I)^j v = 0$ , for all  $j$ . This is a subspace. Indeed, for  $v_1, v_2$  being generalized eigenvectors, then

$$0 + 0 = (T - \lambda I)^k v_1 + (T - \lambda I)^k v_2 = (T - \lambda I)^n (v_1 + v_2)$$

for sufficiently large  $k$ . This shows that  $V_\lambda$  is closed under addition. Also it's obvious that  $\alpha v$  is a generalized eigenvector for any scalar  $\alpha$  and  $v$ .

Now, let  $v \in V_\lambda$  be a generalized eigenvector belonging to  $\lambda$ . Then

$$(T - \lambda I)^k v = 0$$

for some  $k$ . Note that  $Tv - \lambda v = (T - \lambda I)v$  is also a generalized eigenvector, since

$$(T - \lambda I)^k (T - \lambda I)v = (T - \lambda I)(T - \lambda I)^k v = (T - \lambda I)0 = 0$$

Obviously,  $\lambda v$  is a generalized eigenvector. Finally, this means that

$$Tv = (Tv - \lambda v) + \lambda v \in V_\lambda$$

since this is true for the two summands on the right. We conclude that

$$T(V_\lambda) \subseteq V_\lambda$$

- [b] Consider the two polynomials  $(x - \lambda_1)^k$  and  $(x - \lambda_2)^k$ . They are relatively prime polynomial, thus one can use the Euclidean algorithm to find two polynomials  $p(x), q(x)$  such that

$$\begin{aligned} p(x)(x - \lambda_1)^k + q(x)(x - \lambda_2)^k &= \gcd((x - \lambda_1)^k, (x - \lambda_2)^k) \\ &= 1 \end{aligned}$$

Whereas the coefficients of these polynomials are complex numbers, the variable  $x$  can even be a matrix. Plug in  $T$ , to find that

$$p(T)(T - \lambda_1 I)^k + q(T)(T - \lambda_2 I)^k = I$$

Now, assuming there was a vector  $v \in V_{\lambda_1} \cap V_{\lambda_2}$ , then applying both sides of the equation to  $v$ , we find that

$$0 = v$$

and so  $v$  is the zero vector. We conclude that the sum between any two generalized eigenspaces is direct.

We have proven in the textbook that there are enough generalized eigenvectors to span  $V$ , thus  $V$  is a direct sum of generalized eigenspaces.

## Result

3 of 3

- Note that  $\lambda v$  and  $Tv - \lambda v$  are also generalized eigenvectors. Then their sum,  $Tv$  is also one, showing that the space is invariant.
- Over the complex numbers, there are polynomials  $p, q$  such that  $p(x)(x - \lambda_1)^k + q(x)(x - \lambda_2)^k = 1$ . Take  $k$  large enough, and plug in  $T$ . Apply both sides of the equation to find that for  $v \in V_{\lambda_1} \cap V_{\lambda_2}$ ,  $v = 0$ . Thus the sum is direct.

9. a

Let  $V$  be a finite dimensional vector space. And let  $T : V \rightarrow V$  be a linear operator. Let  $K = \text{Ker}(T)$  and  $W = \text{im}(T)$ .

## Step 2

2 of 5

Suppose  $T^2 = T$ . Then for any  $v \in V$  we have  $T^2v = Tv$  i.e.  $T(T(v)) = T(v)$ . That's for each  $w \in W$  we have  $T(w) = w$ .

Conversely suppose,  $T(w) = w, \forall w \in W$ . Choose  $v \in V$ . Then  $T(v) \in W$ . So that  $T(T(v)) = T(v)$  i.e.  $T^2(v) = T$  since  $v \in V$  is arbitrary.

## Step 3

3 of 5

Suppose  $T^2 = T$ . Then for any  $v \in V$  we have  $v = Tv + (v - Tv)$ . Also,  $T(v - Tv) = Tv - T(T(v)) = Tv - Tv = 0$  i.e.  $V = W + K$ .

Now suppose  $u \in W \cap K$ , then  $u \in W$  i.e.  $u = T(u')$  for some  $u' \in V$ . And  $u \in K$  implies  $Tu = 0$ . Now  $0 = Tu = T(T(u')) = T^2(u') = T(u') = u$ . Hence  $W \cap K = \{0\}$ . So that  $V = W \oplus K$ .

Suppose,  $T^2 = T$ . Then by previous section  $V = \text{im}(T) \oplus \text{ker}(T)$ . Let  $\{w_1, \dots, w_m\}$  be a basis of  $W = \text{im}(T)$  and  $\{k_1, \dots, k_n\}$  be a basis of  $K = \text{ker}(T)$ . Since  $V = W \oplus K$  we have,  $\{w_1, \dots, w_m\} \cup \{k_1, \dots, k_n\}$  is a basis for  $V$ . Now  $T(w_1) = w_1, T(w_2) = w_2, \dots, T(w_m) = w_m$  as projection operator is identity on range. Also  $T(k_1) = 0, T(k_2) = 0, \dots, T(k_n) = 0$ . Hence w.r.t. the basis  $\{w_1, \dots, w_m\} \cup \{k_1, \dots, k_n\}$  the matrix of  $T$  will be

$$M = \begin{bmatrix} I_{m \times m} & 0_{m \times n} \\ 0_{n \times m} & 0_{n \times n} \end{bmatrix}.$$

Now trace of  $T$  is same as trace of  $M$  i.e. trace of  $T$  is  $m = \dim(W) = \text{rank}(T)$ .

## Result

5 of 5

When  $T^2 = T$  we show that  $V = \text{im}(T) \oplus \text{ker}(T)$  and then considering matrix of  $T$  w.r.t. basis of  $\text{im}(T)$  and  $\text{ker}(T)$  we show that rank of  $T$  is same as trace of  $T$ .

10. a

Let  $A$  be a  $m \times n$  matrix over  $\mathbb{R}$  and  $B$  be a  $n \times m$  matrix over  $\mathbb{R}$ . Let  $\lambda \in \mathbb{R} - \{0\}$  be an eigenvalue of  $AB$ . Then for some  $x \in \mathbb{R}^m - \{0\}$  we have  $(AB)x = \lambda x$ . That's  $A(Bx) = \lambda x$ . Since,  $\lambda \neq 0$  and  $x \neq 0$  we have  $Bx \neq 0$ . Hence,  $\lambda Bx = B(\lambda x) = B((AB)x) = (BA)(Bx)$ . Now  $Bx \neq 0$  implies  $Bx$  is an eigenvector of  $BA$  corresponding to eigenvalue  $\lambda$ .

Now let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then,

$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

And,

$$BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence, 0 is not an eigenvalue for  $AB$  but for  $BA$ .

Suppose,  $I_m - AB$  is not invertible  $\implies \det(I_m - AB) = 0 \implies 1$  is an eigenvalue of  $AB$ . Now by result in previous section 1 is also an eigenvalue for  $BA$ . Therefore,  $\det(I_n - BA) = 0$  i.e.  $I_n - BA$  is not invertible. Contrapositively,  $I_n - BA$  is invertible implies  $I_m - AB$  is invertible.

Now suppose,  $I_n - BA$  is not invertible. Then  $\det(I_n - BA) = 0$  i.e. 1 is an eigenvalue of  $BA$ . That's there is  $y \in \mathbb{R}^n - \{0\}$  with  $(BA)y = 1y$  i.e.  $B(Ay) = y$ . Now  $y \neq 0$  implies  $Ay \neq 0$ . So  $Ay = A((BA)y) = (AB)(Ay)$ . Since  $Ay \neq 0$ , we can say 1 is an eigenvalue of  $AB$  with eigenvector  $Ay$ . Therefore,  $\det(I_m - AB) = 0$  i.e.  $I_m - AB$  is not invertible. Contrapositively,  $I_m - AB$  is invertible implies  $I_n - BA$  is invertible.

## Result

3 of 3

Note that, 0 is not an eigenvalue of

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

. But 0 is an eigenvalue of

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

# 5

## Chapter 5

### Section 1

1. a

We have to determine the matrices that represent the following rotations of  $\mathbb{R}^3$  :

- (a) angle  $\theta$ , the axis  $e_2$ ,
- (b) angle  $2\pi/3$ , axis contains the vector  $(1, 1, 1)^t$ ,
- (c) angle  $\pi/2$ , axis contains the vector  $(1, 1, 0)^t$

#### Step 2

2 of 8

→ As any linear operator is determined by its action on basis and rotations are linear operators hence we have to check what it will do on basis.

#### Solution :

We will take counterclockwise as positive direction.

- (a) angle  $\theta$ , the axis  $e_2$ ,

Let  $e_1, e_2, e_3$  be the standard basis of  $\mathbb{R}_3$  then we have to check what this rotation does to these basis.

Then the required rotation fixes  $e_2$  and rotates everything around  $e_2$  by angle  $\theta$ . That is, rotation of  $e_1$  and  $e_3$  takes place in  $X - Z$  plane.

The required rotation is

$$\begin{aligned} e_1 &\mapsto (\cos \theta, 0, -\sin \theta)^t \\ e_2 &\mapsto e_2 = (0, 1, 0)^t \\ e_3 &\mapsto (\sin \theta, 0, \cos \theta)^t \end{aligned}$$

Hence the required rotation matrix is

$$\begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

(b) angle  $2\pi/3$ , axis contains the vector  $(1, 1, 1)^t$ ,

$(1, 1, 1)^t$  is at equal angle from all three standard basis  $e_1, e_2, e_3$  and if we join these standard basis along the path of points of equal distance (distance  $\sqrt{2}$ ) from  $(1, 1, 1)^t$ . Then they form a circle passing through  $e_1, e_2, e_3$  and angle between  $e_1$  and  $e_2$  along this circle (i.e. taking  $(1, 1, 1)^t$  as center) is  $2\pi/3$  (as any two of them same angle between them and total angle swapped by circle is  $2\pi$ ). So

$$\begin{aligned} e_1 &\mapsto e_2 = (0, 1, 0)^t \\ e_2 &\mapsto e_3 = (0, 0, 1)^t \\ e_3 &\mapsto e_1 = (1, 0, 0)^t \end{aligned}$$

Hence the required rotation matrix is

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

(c) angle  $\pi/2$ , axis contains the vector  $(1, 1, 0)^t$

Let this rotation sends

$$\begin{aligned} e_1 &\mapsto x = (x_1, x_2, x_3)^t \\ e_2 &\mapsto y = (y_1, y_2, y_3)^t \\ e_3 &\mapsto z = (z_1, z_2, z_3)^t \end{aligned}$$

Then

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2} = 1 = |y| = \sqrt{y_1^2 + y_2^2 + y_3^2} = |z| = \sqrt{z_1^2 + z_2^2 + z_3^2}$$

and angle between  $e_1$  and  $(1, 1, 0)^t$  is equal to angle between  $x$  and  $(1, 1, 0)^t$ .

Similarly angle between  $e_2$  and  $(1, 1, 0)^t$  is equal to angle between  $y$  and  $(1, 1, 0)^t$  and angle between  $e_3$  and  $(1, 1, 0)^t$  is equal to angle between  $z$  and  $(1, 1, 0)^t$ .

As  $(1, 1, 0)^t$  is in  $X - Y$  plane, it makes an angle of  $\pi/2$  with  $e_3$  hence after it's rotation by  $\pi/2$ , it will land in  $X - Y$  plane, i.e.  $z_3 = 0$ . Then

$$\begin{aligned} e_3^t \cdot (1, 1, 0)^t &= (z_1, z_2, 0) \cdot (1, 1, 0)^t \\ (0, 0, 1) \cdot (1, 1, 0)^t &= z_1 + z_2 \\ 0 &= z_1 + z_2 \\ z_1 &= -z_2 \end{aligned}$$



As  $|z| = \sqrt{z_1^2 + z_2^2 + z_3^2} = 1$ , hence  $z_1 = \frac{1}{\sqrt{2}} = -z_2$  and  $z_3 = 0$ .

Now as angle between  $e_1$  and  $(1, 1, 0)^t$  is equal to angle between  $x$  and  $(1, 1, 0)^t$  hence

$$\begin{aligned} e_1^t \cdot (1, 1, 0)^t &= (x_1, x_2, x_3) \cdot (1, 1, 0)^t \\ (1, 0, 0) \cdot (1, 1, 0)^t &= x_1 + x_2 \\ 1 &= x_1 + x_2 \\ x_1 + x_2 &= 1 \end{aligned} \quad (*)$$

Also as  $e_1$  and  $e_3$  are orthonormal hence so is  $x$  and  $z$  so

$$\begin{aligned} x \cdot z^t &= 0 \\ (x_1, x_2, x_3) \cdot \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)^t &= 0 \\ x_1 - x_2 &= 0 \\ x_1 &= x_2 \end{aligned} \quad (**)$$

(\*) and (\*\*) implies that  $x_1 = x_2 = \frac{1}{2}$

As  $|x| = \sqrt{x_1^2 + x_2^2 + x_3^2} = 1$ ,  $x_3^2 = 1 - x_1^2 - x_2^2 = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}$

Hence  $x_3 = \frac{1}{\sqrt{2}}$  or  $x_3 = -\frac{1}{\sqrt{2}}$  As we rotated  $e_1$  counterclockwise about  $(1, 1, 0)^t$  by  $\pi/2$  so it is not hard to imagine geometrically that its  $Z$  - *coordinate* should be negative hence  $x_3 = -\frac{1}{\sqrt{2}}$

Now as angle between  $e_2$  and  $(1, 1, 0)^t$  is equal to angle between  $y$  and  $(1, 1, 0)^t$  hence

$$\begin{aligned} e_2^t \cdot (1, 1, 0)^t &= (y_1, y_2, y_3) \cdot (1, 1, 0)^t \\ (0, 1, 0) \cdot (1, 1, 0)^t &= y_1 + y_2 \\ 1 &= y_1 + y_2 \\ y_1 + y_2 &= 1 \end{aligned} \quad (+)$$

Also as  $e_2$  and  $e_3$  are orthonormal hence so is  $y$  and  $z$  so

$$\begin{aligned} y \cdot z^t &= 0 \\ (y_1, y_2, y_3) \cdot \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)^t &= 0 \\ y_1 - y_2 &= 0 \\ y_1 &= y_2 \end{aligned} \quad (++)$$

(+) and (++) implies that  $y_1 = y_2 = \frac{1}{2}$

As  $|y| = \sqrt{y_1^2 + y_2^2 + y_3^2} = 1$ ,  $y_3^2 = 1 - y_1^2 - y_2^2 = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}$

Hence  $y_3 = \frac{1}{\sqrt{2}}$  or  $y_3 = -\frac{1}{\sqrt{2}}$  As we rotated  $e_2$  counterclockwise about  $(1, 1, 0)^t$  by  $\pi/2$  so it is not hard to imagine geometrically that its  $Z$  - *coordinate* should be positive hence  $y_3 = \frac{1}{\sqrt{2}}$

$$\begin{aligned} e_1 &\mapsto \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{\sqrt{2}}\right)^t \\ e_2 &\mapsto \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right)^t \\ e_3 &\mapsto \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)^t \end{aligned}$$



Hence the required rotation matrix is

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

## Result

We determined the matrices that represent the following rotations of  $\mathbb{R}^3$  :

- (a) angle  $\theta$ , the axis  $e_2$ ,
- (b) angle  $2\pi/3$ , axis contains the vector  $(1, 1, 1)^t$ ,
- (c) angle  $\pi/2$ , axis contains the vector  $(1, 1, 0)^t$

## 2. a

We have to find the complex eigenvalues of the matrix  $A$  that represents a rotation of  $\mathbb{R}^3$  through the angle  $\theta$  about a pole  $u$

## Step 2

2 of 4

**Euler's Theorem** : The  $3 \times 3$  rotation matrices are the orthogonal  $3 \times 3$  matrices with determinant 1, the elements of the special orthogonal group  $SO_3$ .

**Theorem2** : Let  $M$  be the matrix in  $SO_3$  that represents the rotation  $\rho_{(u,\alpha)}$  with  $\text{spin}(u, \alpha)$ .

- (a) The trace of  $M$  is  $1 + 2 \cos \alpha$
- (b) Let  $B$  be another element of  $SO_3$ , and let  $u' = Bu$ . The conjugate  $M' = BMB^t$  represents the rotation  $\rho_{(u',\alpha)}$  with  $\text{spin}(u', \alpha)$

**Solution :**

Let  $M$  be the matrix of rotation by  $\theta$  around pole  $u$ .

As such a rotation fixes pole  $u$  hence 1 is it's one of the eigenvalue. Let other two eigenvalues be  $x, y$ . Then by

**Theorem2** above,

$$\text{Trace}(M) = 1 + 2 \cos \theta$$

also trace is the sum of all eigenvalues of the matrix. Hence

$$\begin{aligned} x + y + 1 &= 1 + 2 \cos \theta \\ x + y &= 2 \cos \theta \end{aligned} \quad (1)$$

Also by **Euler's Theorem**,

$$\det(M) = 1$$

but  $\det(M)$  is the product of eigenvalues hence

$$x \cdot y = 1 \quad (2)$$

From eq(1) and (2), we have

$$\begin{aligned} x + \frac{1}{x} &= 2 \cos \theta \\ x^2 + 1 &= 2x \cos \theta \\ x^2 - 2x \cos \theta + 1 &= 0 \end{aligned}$$

It's two roots are the other two eigenvalues of  $M$ . Hence

$$x = \cos \theta + \sqrt{\cos^2 \theta - 1}$$

And

$$y = \cos \theta - \sqrt{\cos^2 \theta - 1}$$

## Result

4 of 4

The complex eigenvalues of the matrix  $A$  that represents a rotation of  $\mathbb{R}^3$  through the angle  $\theta$  about a pole  $u$  are  $1, \cos \theta + \sqrt{\cos^2 \theta - 1}$ , and  $\cos \theta - \sqrt{\cos^2 \theta - 1}$

## 3. a

If  $n$  is odd, then the function  $f : O_n \rightarrow SO_n \times \{I, -I\}$ , given by the formula

$$f(A) = \left( \frac{1}{(\det A)^{\frac{1}{n}}} A, (\text{sgn } \det A) I \right)$$

is an isomorphism. Let us prove that.

Since every orthogonal matrix is regular (because its columns make a basis for  $\mathbb{C}^n$ , so the matrix is of maximal rank), it follows that  $\det A \neq 0$ .

Since  $n$  is odd,  $x^{\frac{1}{n}}$  is defined for every  $x \in \mathbb{C} \setminus 0$  and  $(x^{\frac{1}{n}})^n = x$ . Therefore,

$$\det\left(\frac{1}{(\det A)^{\frac{1}{n}}} A\right) = \left(\frac{1}{(\det A)^{\frac{1}{n}}}\right)^n \det A = \frac{1}{\det A} \det A = 1 \Rightarrow A \in SO_n$$

This proves that  $f$  is (well) defined.

Let us now prove that  $f$  is a homomorphism.

For  $A, B \in O_n$ , using the **Binet-Cauchy theorem**, we have

$$\begin{aligned} f(AB) &= \left( \frac{1}{(\det(AB))^{\frac{1}{n}}} AB, (\operatorname{sgn} \det(AB))I \right) \\ &= (B \cdot C) = \left( \frac{1}{(\det A)^{\frac{1}{n}} (\det B)^{\frac{1}{n}}} AB, (\operatorname{sgn}(\det A \det B))I^2 \right) \\ &= \left( \frac{1}{(\det A)^{\frac{1}{n}}} A \cdot \frac{1}{(\det B)^{\frac{1}{n}}} B, (\operatorname{sgn} \det A)I \cdot (\operatorname{sgn} \det B)I \right) \\ &= \left( \frac{1}{(\det A)^{\frac{1}{n}}} A, (\operatorname{sgn} \det A)I \right) \cdot \left( \frac{1}{(\det B)^{\frac{1}{n}}} B, (\operatorname{sgn} \det B)I \right) = f(A)f(B) \end{aligned}$$

Let us now prove that  $f$  is a monomorphism.

It is sufficient to prove that  $\ker f = \{I\}$ . We have

$$A \in \ker f \Rightarrow f(A) = (I, I) \Rightarrow \frac{1}{(\det A)^{\frac{1}{n}}} A = I \text{ and } (\operatorname{sgn} \det A)I = I$$

Now  $\frac{1}{(\det A)^{\frac{1}{n}}} A = I \Rightarrow A = (\det A)^{\frac{1}{n}} I$ , which implies that the  $j$ -th column of  $A$  is of form  $(a_{1j}, a_{2j}, \dots, a_{nj})$ , where  $a_{jj} = (\det A)^{\frac{1}{n}}$  and  $a_{ij} = 0$  for  $i \neq j$ . However, since  $A$  is an orthogonal matrix, it must be

$$\|(a_{1j}, a_{2j}, \dots, a_{nj})\| = 1 \Rightarrow (\det A)^{\frac{2}{n}} = 1 \Rightarrow \det A = \pm 1$$

Also,  $(\operatorname{sgn} \det A)I = I$  implies that  $\operatorname{sgn} \det A = 1 \Rightarrow \det A > 0$ , hence  $\det A = 1 \Rightarrow A = I$ .

Now we prove that  $f$  is an epimorphism.

For  $B \in SO_n$  we have that

$$f(B) = \left( \frac{1}{(\det B)^{\frac{1}{n}}} B, (\operatorname{sgn} \det B)I \right) = (\det B = 1 > 0) = (B, I)$$

Also, since  $n$  is odd and  $\det(-B) = -\det B$ , we have

$$f(-B) = \left( \frac{1}{-(\det B)^{\frac{1}{n}}} (-B), -(\operatorname{sgn} \det B)I \right) = (B, -I)$$

Therefore,  $f$  is surjective (= an epimorphism).

The groups  $O_n$  and  $SO_n \times \{I, -I\}$  are not isomorphic for (any) even  $n$ .

If  $f : O_n \rightarrow SO_n \times \{I, -I\}$  is an isomorphism, then

$$f|_{\mathcal{Z}(O_n)} : \mathcal{Z}(O_n) \rightarrow \mathcal{Z}(SO_n \times \{I, -I\})$$

is also an isomorphism. However,

$$\mathcal{Z}(O_n) = \mathcal{Z}(SO_n) = \mathcal{Z}(\{I, -I\}) = \{I, -I\}$$

which implies that  $\mathcal{Z}(SO_n \times \{I, -I\}) = \{I, -I\} \times \{I, -I\}$ .

However, the groups  $\{I, -I\}$  and  $\{I, -I\} \times \{I, -I\}$  can't be isomorphic because they don't have the same cardinality (the first group has 2 elements and the second 4), hence there isn't even a bijection between these two groups.

(If  $n$  is odd, then  $-I \notin SO_n$  because  $\det(-I) = -1 \neq 1$ , hence  $\mathcal{Z}(SO_n) = \{I\}$ , but still  $\mathcal{Z}(O_n) = \{I, -I\}$ . Now the group  $\{I, -I\}$  is isomorphic to  $\{I\} \times \{I, -I\}$ .)

## Result

6 of 6

For odd  $n$ , the groups  $O_n$  and  $SO_n \times \{I, -I\}$  are isomorphic (construct an isomorphism) and for even  $n$  they aren't (because their centers are not isomorphic)

4. a

We have to find geometrically the action of an orthogonal  $3 \times 3$  matrix with determinant  $-1$

## Step 2

2 of 3

**Solution :**

→ **Result of problem 3 :** for  $n$  odd,

$$O_n \cong SO_n \times \{I, -I\}$$

As 3 is odd, any orthogonal  $3 \times 3$  matrix with determinant  $-1$  is a product of an element of  $SO_3$  and  $-I$  hence it should geometrically look like a rotation ( $x$  goes to  $x_{(\theta, \rho)}$ , rotation by an angle  $\theta$  about pole  $\rho$ ) composed with a action of  $-I$  which is sending any element to it's symmetrically opposite point about origin (i.e.  $x$  goes to  $-x$ ).

## Result

3 of 3

Geometrically an orthogonal  $3 \times 3$  matrix with determinant  $-1$  is rotation composed with taking a symmetrically opposite point about origin.

5. a

We are given  $A$ , a  $3 \times 3$  orthogonal matrix with  $\det(A) = 1$ , whose angle of rotation is different from 0 or  $\pi$ . Let  $M = A - A^t$ . We have to show the following

(a)  $M$  has rank 2, and a nonzero vector  $X$  in the nullspace of  $M$  is an eigenvector of  $A$  with eigenvalue 1.

(b) An eigenvector explicitly in terms of the entries of the matrix  $A$ .

## Step 2

2 of 5

**Rank-nullity Theorem :** For a given linear transformation  $\phi : V \rightarrow W$ , we have

$$\text{rank}(\phi(V)) + \dim(\text{nullspace}(\phi)) = \dim(V)$$

**Solution :**

Let  $A$  be the rotation matrix  $(u, \theta)$  with  $\theta \neq 0, \pi$ .

$$M = A - A^t$$

(a)  $M$  has rank 2, and a nonzero vector  $X$  in the nullspace of  $M$  is an eigenvector of  $A$  with eigenvalue 1.

It is enough to show that nullspace of  $M$  is one dimensional.

Let  $X \in \mathbb{R}_{3 \times 3}$ ,

$$\begin{aligned} MX &= 0 \\ (A - A^t)X &= 0 \\ AX - A^tX &= 0 \\ AX &= A^tX \\ A^2X &= X \end{aligned} \quad \text{As } AA^t = A^tA = I$$

i.e.  $X$  is an eigenvalue of  $A^2$ .

As eigenvalues of  $A^2$  are square of eigenvalues of  $A$  and by *Exercise 2* of this section, eigenvalues of  $A$  are  $\{1, \cos \theta + \sqrt{\cos^2 \theta - 1}, \cos \theta - \sqrt{\cos^2 \theta - 1}\}$

As  $\theta \neq 0, \pi$ , other two eigenvalues are not 1 hence their square is not 1. So dimension of eigenspace of  $A^2$  with eigenvalue 1 is one dimensional hence dimension of nullspace of  $M$  is one dimensional.

Hence rank of  $M$  is 2.

(b) An eigenvector explicitly in terms of the entries of the matrix  $A$ .

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Then

$$\begin{aligned} A^t &= \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \\ M = A - A^t &= \begin{bmatrix} 0 & a_{12} - a_{21} & a_{13} - a_{31} \\ a_{21} - a_{12} & 0 & a_{23} - a_{32} \\ a_{31} - a_{13} & a_{32} - a_{23} & 0 \end{bmatrix} \end{aligned}$$

Let  $u = a_{12} - a_{21}, v = a_{13} - a_{31}, w = a_{23} - a_{32}$  then  $M$  looks like,

$$M = \begin{bmatrix} 0 & u & v \\ -u & 0 & w \\ -v & -w & 0 \end{bmatrix}$$

Then

$$MX = \begin{bmatrix} 0 & u & v \\ -u & 0 & w \\ -v & -w & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

As  $A$  is a rotation matrix with angle of rotation  $\theta \neq 0, \pi$  and in which case  $A^t$  is a rotation matrix with angle of rotation  $-\theta$  hence  $A \neq A^t$  hence one of the  $u, v, w$  must be non-zero.



**Case I :**  $u \neq 0$

Then

$$MX = \begin{bmatrix} 0 & u & v \\ -u & 0 & w \\ -v & -w & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

gives the solution  $S1 = \{(\frac{w}{u}x, \frac{v}{u}x, x) | x \in \mathbb{R}\}$

Hence one such eigenvector is  $(\frac{w}{u}, \frac{v}{u}, 1)$ .

**Case II :**  $v \neq 0$

The solution  $S2 = \{(-\frac{u}{v}x, x, \frac{w}{v}x) | x \in \mathbb{R}\}$

Hence one such eigenvector is  $(-\frac{u}{v}, 1, \frac{w}{v})$ .

**Case III :**  $w \neq 0$

The solution  $S3 = \{(x, -\frac{v}{w}x, \frac{u}{w}x) | x \in \mathbb{R}\}$

Hence one such eigenvector is  $(1, -\frac{v}{w}, \frac{u}{w})$ .

## Result

5 of

For  $A$ , a  $3 \times 3$  orthogonal matrix with  $\det(A) = 1$ , whose angle of rotation is different from 0 or  $\pi$

(a)  $A - A^t$  has rank 2, and a nonzero vector  $X$  in the nullspace of it is an eigenvector of  $A$  with eigenvalue 1.

(b) We wrote down such an eigenvector explicitly in terms of the entries of the matrix  $A$ .

## Section 2

1. a

Let  $A$  be an invertible matrix. Using the Cayley-Hamilton Theorem, we have to express  $A^{-1}$  in terms of  $A$ ,  $(\det A)^{-1}$ , and the coefficients of the characteristic polynomial. Also we have to verify the expression in the  $2 \times 2$  case.

### Step 2

2 of 6

**Cayley-Hamilton Theorem :** Let  $p(t) = t^n + c_{n-1}t^{n-1} + \dots + c_1t + c_0$  be the characteristic polynomial of an  $n \times n$  complex matrix  $A$ . Then  $p(A) = A^n + c_{n-1}A^{n-1} + \dots + c_1A + c_0I$  is the zero matrix.



**Solution :**

Let  $A$  be an invertible matrix. Let

$$p(t) = t^n + c_{n-1}t^{n-1} + \dots + c_1t + c_0$$

be it's characteristic polynomial. Then by **Cayley-Hamilton Theorem**, we have,

$$p(A) = A^n + c_{n-1}A^{n-1} + \dots + c_1A + c_0I = 0$$

where  $0$  is  $n \times n$  zero matrix.

$$\begin{aligned} 0 &= A^n + c_{n-1}A^{n-1} + \dots + c_1A + c_0I \\ -c_0I &= A^n + c_{n-1}A^{n-1} + \dots + c_1A \\ -c_0A^{-1} &= A^{-1}A^n + c_{n-1}A^{-1}A^{n-1} + \dots + c_1A^{-1}A \text{ multi. both side by } A^{-1} \\ A^{-1} &= -c_0^{-1}(A^{n-1} + c_{n-1}A^{n-2} + \dots + c_1I) \end{aligned}$$

**For  $2 \times 2$  matrices case :**

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

with  $ad - bc \neq 0$  i.e.  $A$  is invertible.

Then it's characteristic polynomial is given by  $\det(A - xI) = 0$ .

$$\begin{aligned} \det(A - xI) &= 0 \\ \det\left(\begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix}\right) &= 0 \\ (a-x)(d-x) - bc &= 0 \\ x^2 - (a+d)x + ad - bc &= 0 \end{aligned}$$

So the characteristic polynomial is

$$x^2 - (a+d)x + ad - bc = 0$$

Hence

$$A^2 - (a+d)A + (ad - bc)I = 0$$

By the formula we obtained,

$$\begin{aligned}
A^{-1} &= -\frac{1}{ad-bc}(A - (a+d)I) \\
A^{-1} &= -\frac{1}{ad-bc}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - (a+d)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\
A^{-1} &= -\frac{1}{ad-bc}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} a+d & 0 \\ 0 & a+d \end{bmatrix}\right) \\
A^{-1} &= -\frac{1}{ad-bc}\begin{bmatrix} -d & b \\ c & -a \end{bmatrix} \\
A^{-1} &= \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}
\end{aligned}$$

And elementary adjoint formula  $A \text{adj}(A) = \text{adj}(A)A = \det(A)I$  i.e. inverse of  $A$  is  $\frac{1}{\det(A)}\text{adj}(A)$

$$\text{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

So by this formula

$$\begin{aligned}
A^{-1} &= \frac{1}{\det(A)}\text{adj}(A) \\
&= \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}
\end{aligned}$$

Hence the formula we derived using **Cayley-Hamilton Theorem** varifies for  $2 \times 2$  invertible matrices.

## Result

6 of 6

Given  $A$  be an invertible matrix with characteristic polynomial  $p(t) = t^n + c_{n-1}t^{n-1} + \dots + c_1t + c_0$ , the formula for  $A^{-1}$  is

$$A^{-1} = -c_0^{-1}(A^{n-1} + c_{n-1}A^{n-2} + \dots + c_1I)$$

## 2. a

We are given an  $m \times m$  complex matrices  $A$  and an  $n \times n$  complex matrix  $B$  and let linear operator  $T$  on the space  $\mathbb{C}^{m \times n}$  of all complex matrices defined by  $T(M) = AMB$ .

(a) We have to construct an eigenvector for  $T$  out of a pair of column vectors  $X, Y$ , where  $X$  is an eigenvector for  $A$  and  $Y$  is an eigenvector for  $B^t$ .

(b) We have to determine the eigenvalues of  $T$  in terms of those of  $A$  and  $B$ .

(c) We have to determine the trace of this operator.

**Solution :**

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

is  $m \times m$  complex matrices and

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

is  $n \times n$  complex matrix.

(a) We have to construct an eigenvector for  $T$  out of a pair of column vectors  $X, Y$ , where  $X$  is an eigenvector for  $A$  and  $Y$  is an eigenvector for  $B^t$ .

Let  $X = (x_1, x_2, \dots, x_m)^t$  be an eigenvector of  $A$  with eigenvalue  $\lambda_A$  and  $Y = (y_1, y_2, \dots, y_n)^t$  be an eigenvector of  $B^t$  with eigenvalue  $\lambda_{B^t}$ .

Consider the  $m \times n$  matrix

$$M = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \cdots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix}$$

Claim is that  $M$  is the eigenvector of  $T$  with eigenvalue  $\lambda_A \lambda_B$ .

$$\begin{aligned} TM &= AMB \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \cdots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \right) \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \\ &= \left( \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \right) \left( \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}^t \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right)^t \\ &= (AX)(B^t Y)^t \\ &= (\lambda_A X)(\lambda_{B^t} Y)^t \\ &= \lambda_A \lambda_{B^t} \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \right) \\ &= \lambda_A \lambda_{B^t} \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \cdots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} \\ &= \lambda_A \lambda_{B^t} M \end{aligned}$$

Hence the claim.

(b) Determine the eigenvalues of  $T$  in terms of those of  $A$  and  $B$ .

$T$  can have at most  $mn$  eigenvalues. Let  $\lambda_1, \lambda_2, \dots, \lambda_m$  be the eigenvalues of  $A$  and  $\alpha_1, \alpha_2, \dots, \alpha_n$  be eigenvalues of  $B$ . Then as eigenvalues of any matrix and its transpose are the same,  $\alpha_1, \alpha_2, \dots, \alpha_n$  are the eigenvalues of  $B^t$ .

So by part(a) we have  $\lambda_1\alpha_1, \lambda_2\alpha_1, \dots, \lambda_m\alpha_1, \lambda_1\alpha_2, \lambda_2\alpha_2, \dots, \lambda_m\alpha_2, \dots, \lambda_1\alpha_n, \lambda_2\alpha_n, \dots, \lambda_m\alpha_n$  as  $mn$  eigenvalues of  $T$ .

(c) Determine the trace of this operator.

As trace of a linear operator (equivalent to a matrix) is the sum of all eigenvalues hence

$$\begin{aligned} \text{tr}(T) &= \text{sum of all eigenvalues of } T \\ &= \sum_{i=1}^m \sum_{j=1}^n \lambda_i \alpha_j \\ &= \left( \sum_{i=1}^m \lambda_i \right) \left( \sum_{j=1}^n \alpha_j \right) \\ &= \text{tr}(A) \text{tr}(B) \end{aligned}$$

Hence

$$\text{tr}(T) = \text{tr}(A) \text{tr}(B)$$

## Result

6 of 6

Given complex matrices  $A, m \times m$  and  $B, n \times n$  and linear operator  $T$  on the space  $\mathbb{C}^{m \times n}$  of all complex matrices defined by  $T(M) = AMB$ . Then

(a) If  $X$  is an eigenvector of  $A$  and  $Y$  of  $B^t$  then  $XY^t$  is the eigenvector of  $T$ .

(b) If  $\lambda_A$  is an eigenvalue of  $A$  and  $\lambda_B$  of  $B^t$  then  $\lambda_A \lambda_B$  is the eigenvalue of  $T$ .

$$(c) \text{tr}(T) = \text{tr}(A) \text{tr}(B)$$

3. a

Let  $A$  be an  $n \times n$  complex matrix. Consider the operator  $T$  defined on  $n \times n$  complex matrices by

$$T(M) = AM - MA$$

We will use a continuity argument to determine its rank.

First, assume that  $A = D$ , a diagonal matrix with diagonal entries  $\lambda_1, \dots, \lambda_n$ . Let  $E_{ij}$  be the matrix with 1 as the  $(i, j)$ -th entry and 0 for the rest. It is immediate that

$$\begin{aligned} T(E_{ii}) &= DE_{ii} - E_{ii}D \\ &= \lambda_i E_{ii} - \lambda_i E_{ii} \\ &= 0 \end{aligned}$$

Note that  $E_{ij}$  are linearly independent. Thus the kernel contains at least  $n$  linearly independent matrices (the  $E_{ii}$ ). This shows that the rank of  $T$  is at most  $n^2 - n$ .

We see that  $E_{ii}$  are the eigenvectors for 0. For the other eigenvectors, we find:

$$\begin{aligned} T(E_{ij}) &= DE_{ij} - E_{ij}D \\ &= \lambda_i E_{ij} - \lambda_j E_{ij} \\ &= (\lambda_i - \lambda_j) E_{ij} \end{aligned}$$

Thus all other matrices  $E_{ij}$  are also eigenmatrices for  $T$  (and since there are  $n^2$  of them in total, these must be all the eigenmatrices). The eigenvalues are  $\lambda_i - \lambda_j$ .

Now, assume that  $A$  is diagonalizable. Then there is a matrix  $P$  such that

$$P^{-1}AP = D$$

for a diagonal matrix  $D$ . Equivalently, this can be written as  $AP = PD$  and  $P^{-1}A = DP^{-1}$ . Plug in  $PE_{ij}P^{-1}$  into  $T$ , to find

$$\begin{aligned} T(P^{-1}E_{ij}P) &= APE_{ij}P^{-1} - PE_{ij}P^{-1}A \\ &= PDE_{ij}P^{-1} - PE_{ij}DP^{-1} \\ &= P(DE_{ij} - E_{ij}D)P^{-1} \\ &= P(\lambda_i E_{ij} - \lambda_j E_{ij})P^{-1} \\ &= (\lambda_i - \lambda_j)PE_{ij}P^{-1} \end{aligned}$$

This shows that the matrix  $D$  has eigenvalues  $\lambda_i - \lambda_j$  for the eigenmatrices  $PE_{ij}P^{-1}$ . Taking  $i = j$ , we again find that there are at least  $n$  linearly independent matrices in the kernel and so again the rank of  $T$  is at most  $n^2 - n$ .

Finally, let  $A$  be an arbitrary complex matrix. Then choose a sequence of matrices  $A_k$  with distinct eigenvalues which converge to  $A$ . If the eigenvalues of  $A_k$  are  $\lambda_{i,k}$ , then the eigenvalues of

$$T_k(M) = A_k M - M A_k$$

are  $\lambda_{i,k} - \lambda_{j,k}$  just as above. Now, we simply take the limit  $k \rightarrow \infty$ . Note that the eigenvalues  $\lambda_{i,k}$  converge to  $\lambda_i$ . Since all the matrices  $A_k$  have rank less than  $n^2 - n$ , they have at least  $n$  eigenvalues 0. Thus, at least  $n$  of the  $\lambda_{i,k}$  are always zero for a given  $k$ . Reordering the eigenvalues of  $A_k$  if necessary, we find that  $\lambda_{i,k} \rightarrow 0$  for at least  $n$  eigenvalues. Thus  $A$  will have again  $n$  times the eigenvalue 0. We conclude that the rank of  $A$  is at most  $n^2 - n$ .

Note that the same logic applies to the rest of the eigenvalues. We have

$$\lambda_{i,k} - \lambda_{j,k} \xrightarrow{k \rightarrow \infty} \lambda_i - \lambda_j$$

Thus the eigenvalues of  $T$  are given by  $\lambda_i - \lambda_j$  (the differences of the eigenvalues of  $A$ ).

The exercise is solved by a continuity argument.  
The statement is first checked for diagonalizable matrices.

There is always a sequence of diagonalizable matrices that converges to an arbitrary matrix. Taking such a sequence, the result follows.

4. a

Given  $A$  and  $B$ , diagonalizable complex matrices, we have to prove that there is an invertible matrix  $P$  such that  $P^{-1}AP$  and  $P^{-1}BP$  are both diagonal if and only if  $AB = BA$ .

**Jordan Decomposition Theorem(Matrix form)** : Let  $A$  be an  $n \times n$  complex matrix. There is an invertible complex matrix  $P$  such that  $P^{-1}AP$  has Jordan form. Where Jordan form is a block matrix where each block correspond to an eigenvalue and no two blocks correspond to same eigenvalue.

→ **Diagonal matrices of same order commute with each other** Let

$$D_1 = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{12} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{1n} \end{bmatrix}$$

$$D_2 = \begin{bmatrix} d_{21} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{2n} \end{bmatrix}$$

be two diagonal matrices. Then

$$\begin{aligned} D_1 D_2 &= \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{12} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{1n} \end{bmatrix} \begin{bmatrix} d_{21} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{2n} \end{bmatrix} \\ &= \begin{bmatrix} d_{11}d_{21} & 0 & \cdots & 0 \\ 0 & d_{12}d_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{1n}d_{2n} \end{bmatrix} \\ &= \begin{bmatrix} d_{21}d_{11} & 0 & \cdots & 0 \\ 0 & d_{22}d_{12} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{2n}d_{1n} \end{bmatrix} \\ &= \begin{bmatrix} d_{21} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{2n} \end{bmatrix} \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{12} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_{1n} \end{bmatrix} \\ &= D_2 D_1 \end{aligned}$$



**Proof :**

We will take all matrices to be  $n \times n$ .

( $\implies$ ) Let there is an invertible matrix  $P$  such that  $P^{-1}AP$  and  $P^{-1}BP$  are both diagonal.

Then as  $P^{-1}AP$  and  $P^{-1}BP$  are diagonal,

$$\begin{aligned} P^{-1}APP^{-1}BP &= P^{-1}BPP^{-1}AP \\ P^{-1}ABP &= P^{-1}BAP && \text{multi. both side} \\ AB &= BA && \text{in right by } P^{-1} \text{ and left by } P \end{aligned}$$

Hence  $A$  and  $B$  commute.

( $\impliedby$ ) Let  $A$  and  $B$  commute *i.e.*  $AB = BA$ .

As  $A$  is diagonalizable, let  $\lambda_1, \lambda_2, \dots, \lambda_r$  be distinct eigenvalues of  $A$ . Then by **Jordan Decomposition Theorem**, there exist an invertible matrix  $P'$  such that  $P'^{-1}AP'$  is block matrix and each it looks like

$$P'^{-1}AP' = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & A_r \end{bmatrix}$$

where  $A_i$  is eigenspace corresponding to  $\lambda_i$ . But as  $A$  is diagonalizable, we can assume that  $P'^{-1}AP'$  is diagonal *i.e.* each  $A_i$  are diagonal *i.e.*

$$A_i = \begin{bmatrix} \lambda_i & 0 & \dots & 0 \\ 0 & \lambda_i & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_i \end{bmatrix}$$

Now given any eigenvector  $X$  of  $A$  with eigenvalue  $\lambda$ ,

$$A(BX) = B(AX) = \lambda BX$$

hence  $BX$  belong to the eigenspace of eigenvalue  $\lambda$ . Hence  $P'^{-1}BP'$  looks like

$$P'^{-1}BP' = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & B_r \end{bmatrix}$$

With  $B_i$  being block matrix corresponding to  $\lambda_i$ . Now as  $B$  is diagonalizable, each  $B_i$  is diagonalizable. Let  $R_i$  be the invertible matrices such that  $R_i^{-1}B_iR_i$  is diagonal for all  $1 \leq i \leq r$ .

Now consider

$$R = \begin{bmatrix} R_1 & 0 & \cdots & 0 \\ 0 & R_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r \end{bmatrix}$$

$$R^{-1}P'^{-1}BP'R = \begin{bmatrix} R_1^{-1} & 0 & \cdots & 0 \\ 0 & R_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r^{-1} \end{bmatrix} \begin{bmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & B_r \end{bmatrix} \begin{bmatrix} R_1 & 0 & \cdots & 0 \\ 0 & R_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r \end{bmatrix}$$

$$= \begin{bmatrix} R_1^{-1}B_1R_1 & 0 & \cdots & 0 \\ 0 & R_2^{-1}B_2R_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r^{-1}B_rR_r \end{bmatrix}$$

Which is diagonal.

As each  $A_i$  is  $\lambda_i I$  hence commute.

$$R^{-1}P'^{-1}AP'R = \begin{bmatrix} R_1^{-1} & 0 & \cdots & 0 \\ 0 & R_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_r \end{bmatrix} \begin{bmatrix} R_1 & 0 & \cdots & 0 \\ 0 & R_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r \end{bmatrix}$$

$$= \begin{bmatrix} R_1^{-1}\lambda_1 I R_1 & 0 & \cdots & 0 \\ 0 & R_2^{-1}\lambda_2 I R_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & R_r^{-1}\lambda_r I R_r \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 R_1^{-1}R_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 R_2^{-1}R_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_r R_r^{-1}R_r \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_r I \end{bmatrix}$$

Hence  $P = P'R$  is the required matrix which diagonalizes both  $A$  and  $B$  simultaneously.

## Result

6 of

Let  $A$  and  $B$ , diagonalizable complex matrices. Then there is an invertible matrix  $P$  such that  $P^{-1}AP$  and  $P^{-1}BP$  are both diagonal if and only if  $AB = BA$ .

## Section 3

1. a

We have to prove the product rule for differentiation of matrix-valued functions *i.e.* given two  $n \times n$  matrix-valued functions  $A(x), B(x)$ ,  $\frac{d(A(x)B(x))}{dx} = \frac{dA(x)}{dx}B(x) + A(x)\frac{dB(x)}{dx}$

## Step 2

2 of

→ **Product rule for scalar functions** : Let  $f(x), g(x)$  be two scalar real-valued function. Then

$$\frac{d(fg)}{dx} = \frac{df}{dx}g + f\frac{dg}{dx}$$

**Proof :**

We will write  $A(x)$  as  $A$  and  $B(x)$  as  $B$ .

Let  $A = ((a_{ij}))_{1 \leq i, j \leq n}$  *i.e.*  $A$ 's  $(i, j)$ th entry is  $a_{ij}$ . Similarly  $B = ((b_{ij}))_{1 \leq i, j \leq n}$ . Then their product is

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \cdots & \sum_{k=1}^n a_{1k}b_{kn} \\ \sum_{k=1}^n a_{2k}b_{k1} & \sum_{k=1}^n a_{2k}b_{k2} & \cdots & \sum_{k=1}^n a_{2k}b_{kn} \\ \vdots & \vdots & \cdots & \vdots \\ \sum_{k=1}^n a_{nk}b_{k1} & \sum_{k=1}^n a_{nk}b_{k2} & \cdots & \sum_{k=1}^n a_{nk}b_{kn} \end{bmatrix} \end{aligned}$$

Hence

$$AB = ((\sum_{k=1}^n a_{ik}b_{kj}))_{1 \leq i, j \leq n}$$

So

$$\begin{aligned} \frac{d(AB)}{dx} &= \frac{d((\sum_{k=1}^n a_{ik}b_{kj}))_{1 \leq i, j \leq n}}{dx} \\ &= ((\sum_{k=1}^n \frac{d(a_{ik}b_{kj})}{dx}))_{1 \leq i, j \leq n} \\ &= ((\sum_{k=1}^n \frac{da_{ik}}{dx} b_{kj} + a_{ik} \frac{db_{kj}}{dx}))_{1 \leq i, j \leq n} \\ &= ((\sum_{k=1}^n \frac{da_{ik}}{dx} b_{kj}))_{1 \leq i, j \leq n} + ((\sum_{k=1}^n a_{ik} \frac{db_{kj}}{dx}))_{1 \leq i, j \leq n} \\ &= \frac{dA}{dx}B + A\frac{dB}{dx} \end{aligned}$$

Hence the product rule.

## Result

Given two  $n \times n$  matrix-valued functions  $A(x), B(x)$ ,

$$\frac{d(A(x)B(x))}{dx} = \frac{dA(x)}{dx}B(x) + A(x)\frac{dB(x)}{dx}$$

2. a

$A(t)$  and  $B(t)$  are differentiable matrix-valued functions on  $t$ . We have to compute the derivatives of the following :

(a)  $A(t)^3$

(b)  $A(t)^{-1}$

(c)  $A(t)^{-1}B(t)$

## Step 2

2 of 6

→ **Product rule for differentiation of matrix-valued functions** : Let  $A_1, \dots, A_k$  be differentiable matrix-valued functions of  $t$ , of suitable sizes so that their product is defined. Then the matrix product  $A_1 \cdots A_k$  is differentiable, and its derivative is

$$\frac{d}{dt} (A_1 \cdots A_k) = \sum_{i=1}^k A_1 \cdots A_{i-1} \left( \frac{dA_i}{dt} \right) A_{i+1} \cdots A_k$$

### Solution :

We will denote any matrix-valued function  $M(t)$  over  $t$  as  $M$ .

(a)  $A(t)^3$

$$\begin{aligned} \frac{d(A^3)}{dt} &= \frac{d(AAA)}{dt} \\ &= \frac{dA}{dt}AA + A\frac{dA}{dt}A + AA\frac{dA}{dt} \text{ by diff. prod. rule} \end{aligned}$$

Hence

$$\frac{d(A(t)^3)}{dt} = \frac{d(A(t))}{dt}A(t)A(t) + A(t)\frac{d(A(t))}{dt}A(t) + A(t)A(t)\frac{dA(t)}{dt}$$

(b)  $A(t)^{-1}$

$$\begin{aligned} AA^{-1} &= I \\ \frac{d(AA^{-1})}{dt} &= \frac{dI}{dt} \\ \frac{dA}{dt}A^{-1} + A\frac{d(A^{-1})}{dt} &= 0 \\ A\frac{d(A^{-1})}{dt} &= -\frac{dA}{dt}A^{-1} \\ \frac{d(A^{-1})}{dt} &= -A\frac{dA}{dt}A^{-1} \end{aligned}$$

Hence

$$\frac{d(A(t)^{-1})}{dt} = -A(t)\frac{dA(t)}{dt}A(t)^{-1}$$

(c)  $A(t)^{-1}B(t)$

$$\begin{aligned}\frac{d(A^{-1}B)}{dt} &= \frac{A^{-1}}{dt}B + A^{-1}\frac{dB}{dt} \\ &= A\frac{dA}{dt}A^{-1}B + A^{-1}\frac{dB}{dt} \quad \text{by (b)}\end{aligned}$$

Hence

$$\frac{d(A(t)^{-1}B(t))}{dt} = A(t)\frac{dA(t)}{dt}A(t)^{-1}B(t) + A(t)^{-1}\frac{dB(t)}{dt}$$

## Result

$A(t)$  and  $B(t)$  are differentiable matrix-valued functions on  $t$ . Then

$$\begin{aligned}\frac{d(A(t)^3)}{dt} &= \frac{d(A(t))}{dt}A(t)A(t) + A(t)\frac{d(A(t))}{dt}A(t) + A(t)A(t)\frac{d(A(t))}{dt} \\ \frac{d(A(t)^{-1})}{dt} &= A(t)\frac{dA(t)}{dt}A(t)^{-1} \\ \frac{d(A(t)^{-1}B(t))}{dt} &= A(t)\frac{dA(t)}{dt}A(t)^{-1}B(t) + A(t)^{-1}\frac{dB(t)}{dt}\end{aligned}$$

3. a

We have to solve the equation  $\frac{dX}{dt} = AX$  for the following matrices :

(a)

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$

(c)

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & -1 \end{bmatrix}$$

(d)

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

**Theorem :** Let  $A$  be an  $n \times n$  matrix, and let  $P$  be an invertible matrix such that  $\Lambda = P^{-1}AP$  is diagonal, with diagonal entries  $\lambda_1, \dots, \lambda_n$ . The general solution of the system  $\frac{dX}{dt} = AX$  is  $X = P\tilde{X}$ , where  $\tilde{X} = (c_1e^{\lambda_1 t}, \dots, c_ne^{\lambda_n t})^t$  solves the equation  $\frac{d\tilde{X}}{dt} = \Lambda\tilde{X}$ .

**Solution :**

We have to find eigenvectors and corresponding eigenvalues, then using above theorem, we can easily solve the equation.

(a)

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

To find eigenvalues, we have to solve the characteristic equation :

$$\begin{aligned} \det\left(\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}\right) &= 0 \\ (2-\lambda)^2 - 1 &= 0 \\ (\lambda-2-1)(\lambda-2+1) &= 0 \\ (\lambda-3)(\lambda-1) &= 0 \end{aligned}$$

So eigenvalues are 1, 3. And it is easy to check that  $(1, -1)^t$  and  $(1, 1)^t$  are the eigenvectors corresponding to 1 and 3 respectively. So if we take

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

and

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

, then we have,

$$P^{-1} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} P = D$$

Hence the solution is  $X = PX'$  where  $X' = (c_1 e^t, c_2 e^{3t})^t$  and  $c_1, c_2$  arbitrary. So the solution is

$$\begin{aligned} PX' &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} c_1 e^t \\ c_2 e^{3t} \end{bmatrix} \\ &= \begin{bmatrix} c_1 e^t + c_2 e^{3t} \\ -c_1 e^t + c_2 e^{3t} \end{bmatrix} \end{aligned}$$



(b)

$$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$

Characteristic equation :

$$\begin{aligned} \det\left(\begin{bmatrix} 1-\lambda & i \\ -i & 1-\lambda \end{bmatrix}\right) &= 0 \\ (1-\lambda)^2 - 1 &= 0 \\ (\lambda-1-1)(\lambda-1+1) &= 0 \\ \lambda(\lambda-2) &= 0 \end{aligned}$$

So eigenvalues are 0, 2. And it is easy to check that  $(1, i)^t$  and  $(1, -i)^t$  are the eigenvectors corresponding to 0 and 2 respectively. So if we take

$$P = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

and

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

, then we have,

$$P^{-1} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} P = D$$

Hence the solution is  $X = PX'$  where  $X' = (c_1, c_2 e^{2t})^t$  and  $c_1, c_2$  arbitrary. So the solution is

$$\begin{aligned} PX' &= \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 e^{2t} \end{bmatrix} \\ &= \begin{bmatrix} c_1 + c_2 e^{2t} \\ -c_1 i + c_2 e^{2t} i \end{bmatrix} \end{aligned}$$

(c)

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & -1 \end{bmatrix}$$

Characteristic equation :

$$\begin{aligned} \det\left(\begin{bmatrix} 1-\lambda & 2 & 3 \\ 0 & -\lambda & 4 \\ 0 & 0 & -1-\lambda \end{bmatrix}\right) &= 0 \\ (1-\lambda)(-\lambda)(-1-\lambda) &= 0 \\ (1-\lambda)(\lambda)(1+\lambda) &= 0 \end{aligned}$$

So eigenvalues are 0, 1, -1. And it is easy to check that  $(-2, 1, 0)^t$ ,  $(1, 0, 0)^t$  and  $(5, -8, 2)^t$  are the eigenvectors corresponding to 0, 1 and -1 respectively. So if we take

$$P = \begin{bmatrix} -2 & 1 & 5 \\ 1 & 0 & -8 \\ 0 & 0 & 2 \end{bmatrix}$$

and

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

, then we have,

$$P^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & -1 \end{bmatrix} P = D$$

Hence the solution is  $X = PX'$  where  $X' = (c_1, c_2 e^t, c_3 e^{-t})^t$  and  $c_1, c_2, c_3$  arbitrary. So the solution is

$$\begin{aligned} PX' &= \begin{bmatrix} -2 & 1 & 5 \\ 1 & 0 & -8 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 e^t \\ c_3 e^{-t} \end{bmatrix} \\ &= \begin{bmatrix} -2c_1 + c_2 e^t + 5c_3 e^{-t} \\ c_1 - 8c_3 e^{-t} \\ 2c_3 e^{-t} \end{bmatrix} \end{aligned}$$

(d)

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Characteristic equation :

$$\begin{aligned} \det \left( \begin{bmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{bmatrix} \right) &= 0 \\ -\lambda^3 - 1 &= 0 \\ (\lambda - 1)(\lambda^2 + \lambda + 1) &= 0 \end{aligned}$$

So eigenvalues are  $1, \omega, \omega^2$  where  $\omega = \frac{-1 + \sqrt{-3}}{2}$ . And it is easy to check that  $(1, 1, 1)^t, (1, \omega^2, \omega)^t$  and  $(1, \omega, \omega^2)^t$  are the eigenvectors corresponding to  $1, \omega$  and  $\omega^2$  respectively. So if we take

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{bmatrix}$$

and

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{bmatrix}$$

, then we have,

$$P^{-1} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} P = D$$

Hence the solution is  $X = PX'$  where  $X' = (c_1, c_2 e^t, c_3 e^{-t})^t$  and  $c_1, c_2, c_3$  arbitrary. So the solution is

$$\begin{aligned} PX' &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{bmatrix} \begin{bmatrix} c_1 e^t \\ c_2 e^{\omega t} \\ c_3 e^{\omega^2 t} \end{bmatrix} \\ &= \begin{bmatrix} c_1 e^t + c_2 e^{\omega t} + c_3 e^{\omega^2 t} \\ c_1 e^t + c_2 \omega^2 e^{\omega t} + c_3 \omega e^{\omega^2 t} \\ c_1 e^t + c_2 \omega e^{\omega t} + c_3 \omega^2 e^{\omega^2 t} \end{bmatrix} \end{aligned}$$

## Result

7 of 7

Using our knowledge of basic linear algebra and matrix theory, we solved the equation  $\frac{dX}{dt} = AX$  for certain matrices  $A$ .

4. a

Given  $A$  and  $B$  be constant matrices, with  $A$  invertible, we have to solve the inhomogeneous differential equation  $\frac{dX}{dt} = AX + B$  in terms of the solutions to the equation  $\frac{dX}{dt} = AX$ .

**Solution :**

Let  $X_0$  be the solution to the equation  $\frac{dX}{dt} = AX$  and let

$$B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Then the solution to the inhomogeneous differential equation  $\frac{dX}{dt} = AX + B$  is

$$X_0 + \begin{bmatrix} b_1 t \\ b_2 t \\ b_3 t \end{bmatrix}$$

i.e.  $X_0 + Bt$  as

$$\begin{aligned} \frac{d\left(X_0 + \begin{bmatrix} b_1 t \\ b_2 t \\ b_3 t \end{bmatrix}\right)}{dt} &= \frac{dX_0}{dt} + \frac{d\left(\begin{bmatrix} b_1 t \\ b_2 t \\ b_3 t \end{bmatrix}\right)}{dt} \\ &= AX_0 + \begin{bmatrix} \frac{d(b_1 t)}{dt} \\ \frac{d(b_2 t)}{dt} \\ \frac{d(b_3 t)}{dt} \end{bmatrix} \\ &= AX_0 + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\ &= AX_0 + B \end{aligned}$$

**Result**

3 of 3

If  $X_0$  is the solution of the equation  $\frac{dX}{dt} = AX$ , then the solution of the equation  $\frac{dX}{dt} = AX + B$  is  $X_0 + Bt$ .

## Section 4

1. a

- [a)] Let  $A$  be the matrix

$$\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$$

. We will calculate the matrix exponential from the definition. It is not hard to see that

$$A^n = \begin{bmatrix} a^n & ba^{n-1} \\ 0 & 0 \end{bmatrix}$$

Using this, we calculate the matrix exponential with the exponential series:

$$\begin{aligned} e^A &= I + \sum_{n=1}^{\infty} \frac{A^n}{n!} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sum_{n=1}^{\infty} \frac{1}{n!} \begin{bmatrix} a^n & ba^{n-1} \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 + \sum_{n=1}^{\infty} \frac{a^n}{n!} & \sum_{n=1}^{\infty} \frac{ba^{n-1}}{n!} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} e^a & \frac{b}{a} \sum_{n=1}^{\infty} \frac{a^n}{n!} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} e^a & \frac{b}{a}(e^a - 1) \\ 0 & 1 \end{bmatrix} \end{aligned}$$

- [b)] We will use the definition again. Notice that the power of this matrix is:

$$\begin{aligned} A^n &= \begin{bmatrix} 2\pi i & 2\pi i \\ 0 & 2\pi i \end{bmatrix}^n \\ &= (2\pi i)^n \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}^n \\ &= (2\pi i)^n \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}^n \end{aligned}$$

We calculate the exponential:

$$\begin{aligned} e^A &= I + \sum_{n=1}^{\infty} \frac{A^n}{n!} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sum_{n=1}^{\infty} \frac{(2\pi i)^n}{n!} \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 + \sum_{n=1}^{\infty} \frac{(2\pi i)^n}{n!} & \sum_{n=1}^{\infty} \frac{n(2\pi i)^n}{n!} \\ 0 & 1 + \sum_{n=1}^{\infty} \frac{(2\pi i)^n}{n!} \end{bmatrix} \\ &= \begin{bmatrix} e^{2\pi i} & (2\pi i) \sum_{n=0}^{\infty} \frac{(2\pi i)^n}{n!} \\ 0 & e^{2\pi i} \end{bmatrix} \\ &= \begin{bmatrix} e^{2\pi i} & (2\pi i)e^{2\pi i} \\ 0 & e^{2\pi i} \end{bmatrix} \end{aligned}$$

Now, simply recall that  $e^{2\pi i} = 1$ , so

$$e^A = \begin{bmatrix} 1 & 2\pi i \\ 0 & 1 \end{bmatrix}$$

- [c)] We will use the definition again. The power of this matrix is not hard to calculate, but it will vary depending on parity:

$$\begin{aligned}
 A^2 &= \begin{bmatrix} b^2 & 0 \\ 0 & b^2 \end{bmatrix} \\
 &= -b^2 I \\
 A^n &= \begin{cases} (-1)^{\frac{n}{2}} b^n I & n \text{ even} \\ (-1)^{\frac{n-1}{2}} b^{n-1} A & n \text{ odd} \end{cases}
 \end{aligned}$$

We calculate the exponential:

$$\begin{aligned}
 e^A &= \sum_{n=0}^{\infty} \frac{A^n}{n!} \\
 &= \sum_{n \text{ even}} (-1)^{\frac{n}{2}} \frac{b^n}{n!} I + \sum_{n \text{ odd}} (-1)^{\frac{n-1}{2}} \frac{b^{n-1}}{n!} A
 \end{aligned}$$

Note that we only need to sum the coefficients of the matrices. Let  $n = 2k$  in the first sum and  $n = 2k + 1$  in the second. Then

$$\begin{aligned}
 e^A &= \sum_{k=0}^{\infty} (-1)^k \frac{b^{2k}}{(2k)!} I + \sum_{k=0}^{\infty} (-1)^k \frac{b^{2k}}{(2k+1)!} A \\
 &= (\cos b) I + (\sin b) \frac{A}{b} \\
 &= \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}
 \end{aligned}$$

- [d)] We will use the definition again. The power of this matrix is similar to part b). One can also use here that the matrix is given in Jordan form. Write  $A = I + N$ , where  $N$  is nilpotent. Then  $I$  and  $N$  commute so

$$\begin{aligned}
 e^A &= e^{I+N} = e^I e^N \\
 &= e^I \left( I + \frac{N}{1!} \right) \\
 &= \begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix} \left( I + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right) \\
 &= \begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} e & 0 \\ e & e \end{bmatrix}
 \end{aligned}$$

- [e)] We will use the definition again. The matrix is nilpotent and in fact we have

$$A^2 = \begin{bmatrix} 0 & & \\ 0 & 0 & \\ 1 & 0 & 0 \end{bmatrix}$$

Then

$$\begin{aligned}
 e^A &= I + A + \frac{A^2}{2!} \\
 &= \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \\ 1 & 1 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1/2 & 1 & 1 \end{bmatrix}
 \end{aligned}$$

## Result

In each of a),b),c),d),e) we compute the exponential by the exponential series.

### 2. a

Let  $A$  be a matrix over the complex numbers. Recall that  $Pe^AP^{-1} = e^{PAP^{-1}}$ . Consider the determinant:

$$\begin{aligned}\det(e^A) &= \det(P^{-1}) \det(e^A) \det(P) \\ &= \det(P^{-1}e^AP) \\ &= \det(e^{P^{-1}AP}) \\ &= \det(e^J)\end{aligned}$$

$J$  here is the Jordan normal form. This can be expressed as a sum of a diagonal and nilpotent matrix which commute. Suppose  $J = D + N$  is this decomposition. Then

$$\begin{aligned}e^J &= e^{D+N} = e^D e^N \\ &= e^D \left( I + N + \frac{N^2}{2} + \dots \right)\end{aligned}$$

This is a lower triangular matrix. The nilpotent part doesn't play a role on the diagonal. Thus the determinant of  $e^J$  is equal to the determinant of  $e^D$ .  $e^D$  is calculated by exponentiating all diagonal entries, thus we finally have that

$$\det(e^A) = \det(e^J) = \prod_{i=1}^n e^{\lambda_i}$$

where the  $\lambda_i$  are all the eigenvalues of  $A$  (with multiplicity). Now, recall that the trace can be simply calculated as the sum of the eigenvalues. Thus

$$\begin{aligned}\det(e^A) &= \prod_{i=1}^n e^{\lambda_i} \\ &= e^{\sum_{i=1}^n \lambda_i} \\ &= e^{\text{trace } A}\end{aligned}$$

## Result

2 of 2

We have that  $\det(e^A) = \det(e^J)$  where  $J$  is the Jordan form of  $A$ . Only the diagonal of  $J$  contributes to the diagonal of  $e^J$ , and so  $\det(e^A) = \det(e^J) = \prod e^{\lambda} = e^{\sum \lambda} = e^{\text{trace } A}$ .

### 3. a

- [a] Let  $X$  be an eigenvector of an  $n \times n$  matrix  $A$ , with eigenvalue  $k$ . Assume that  $A$  is invertible, then

$$\begin{aligned}AX &= kX && \bigg/ A^{-1} \\ (A^{-1}A)X &= A^{-1}(kX) && \bigg/ \cdot k^{-1} \\ A^{-1}X &= k^{-1}X\end{aligned}$$

We see that  $k^{-1}$  is an eigenvalue of  $A^{-1}$ .



- [b]) Let  $X$  be an eigenvector of an  $n \times n$  matrix  $A$ , with eigenvalue  $k$ . Note that  $A^k X = \lambda^k X$ . Then

$$e^A X = \left( \sum_{n=0}^{\infty} \frac{A^n}{n!} \right) X$$

Now, matrix multiplication is a continuous operation, thus  $X$  can go under the limit to get:

$$\begin{aligned} e^A X &= \left( \sum_{n=0}^{\infty} \frac{A^n}{n!} X \right) \\ &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} X \\ &= e^\lambda X \end{aligned}$$

We see that  $X$  is an eigenvector for the eigenvalue  $e^\lambda$ .

### Result

- Multiply  $AX = kx$  by  $(kA)^{-1}$ .
- Note that  $A^k X = \lambda^k X$ . The statement follows from the definition of the matrix exponential.

### 4. a

Let  $A$  and  $B$  be commuting matrices. We will prove  $e^{A+B} = e^A e^B$  by expanding the exponential series. One has

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{\infty} \frac{(A+B)^n}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{n!} \binom{n}{k} A^k B^{n-k} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{k!(n-k)!} A^k B^{n-k} \\ e^A e^B &= \left( \sum_{i=0}^{\infty} \frac{A^i}{i!} \right) \left( \sum_{j=0}^{\infty} \frac{B^j}{j!} \right) \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{A^k}{k!} \frac{B^{n-k}}{(n-k)!} \end{aligned}$$

Note that  $e^A e^B$  is a product of series and so for the infinite product to be convergent, one of the series should be absolutely convergent. The exponential series certainly satisfies this. Thus we see that both expressions are equal and so  $e^{A+B} = e^A e^B$ .

### Result

2 of 2

This is an easy application of the binomial theorem and product of infinite series. Note that a product of two series converges if one of the series converges absolutely.

### 5. a

- [a)] We solve the differential equation  $\frac{dX}{dt} = AX$  where  $A$  is the matrix given by

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$$

Note that the matrix is in Jordan normal form. The solution of the differential equation is given explicitly by  $e^{tA}$  so we compute it. Let  $D$  be the diagonal part and  $N$  the nilpotent part (note that  $N^2 = 0$ ). Then

$$\begin{aligned} e^{tA} &= e^t(D + N) \\ &= e^{tD}e^{tN} \\ &= e^{tD}(I + tN) \\ &= \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix} \\ &= \begin{bmatrix} e^{2t} & 0 \\ te^{2t} & e^{2t} \end{bmatrix} \end{aligned}$$

Taking  $X_1$  to be the first column of this matrix and  $X_2$  to be the second, gives two possible linearly independent solutions (which span the solution space).

- [b)] We solve the differential equation  $\frac{dX}{dt} = AX$  where  $A$  is the matrix given by

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Note that the matrix is in Jordan normal form. The solution of the differential equation is given explicitly by  $e^{tA}$  so we compute it. The matrix is already nilpotent. One has

$$\begin{aligned} e^{tA} &= I + tA \\ &= \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix} \end{aligned}$$

The solution space is spanned by the columns of this matrix.

- [c)] We solve the differential equation  $\frac{dX}{dt} = AX$  where  $A$  is the matrix given by

$$A = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ & 1 & 1 \end{bmatrix}$$

Note that the matrix is in Jordan normal form. The solution of the differential equation is given explicitly by  $e^{tA}$  so we compute it. Let  $D$  be the diagonal part and  $N$  the nilpotent part (note that  $N^3 = 0$ ). Then

$$\begin{aligned} e^{tA} &= e^t(D + N) \\ &= e^{tD}e^{tN} \\ &= e^{tD}\left(I + tN + \frac{t^2}{2}N^2\right) \\ &= \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^t \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ t^2/2 & t & 1 \end{bmatrix} \\ &= \begin{bmatrix} e^t & 0 & 0 \\ te^t & e^t & 0 \\ t^2/2e^t & te^t & e^t \end{bmatrix} \end{aligned}$$

The columns give the generators for space of solutions.

## Result

4 of 4

In all of the cases the matrix exponential isn't hard to calculate since the matrices are in Jordan normal form.

6. a

- [a] Let  $A$  be an  $n \times n$  matrix. Define  $\sin A$  and  $\cos A$  be the Taylor series expansions:

$$\begin{aligned}\sin A &= \sum_{k=0}^{\infty} (-1)^k \frac{A^{2k+1}}{(2k+1)!} \\ \cos A &= \sum_{k=0}^{\infty} (-1)^k \frac{A^{2k}}{(2k)!}\end{aligned}$$

One could prove directly that these series converge by considering the matrix norms. But one can also argue as follows:

$$\begin{aligned}\sin A &= \sum_{k=0}^{\infty} (-1)^k \frac{A^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{1}{2i} \left( i^k \frac{A^k}{(k)!} - (-i)^k \frac{A^k}{(k)!} \right) \\ &= \frac{1}{2i} \left( \sum_{k=0}^{\infty} i^k \frac{A^k}{(k)!} - \sum_{k=0}^{\infty} (-i)^k \frac{A^k}{(k)!} \right)\end{aligned}$$

We know that the two series above converge. This means that the series for  $\sin A$ . The same follows for  $\cos A = \frac{1}{2}(e^{iA} + e^{-iA})$ .

- [b] We use the above logic. Write

$$\sin tA = \frac{1}{2i}(e^{t(iA)} - e^{t(-iA)})$$

The right hand side is differentiable and therefore the left hand side as well. Differentiating and using the usual rules we find that

$$\begin{aligned}\frac{d \sin tA}{dt} &= \frac{1}{2i}(iAe^{t(iA)} - (-iA)e^{t(-iA)}) \\ &= \frac{A}{2}(e^{t(iA)} + e^{t(-iA)}) \\ &= A \cos tA\end{aligned}$$

as was to be shown.

## Result

3 of 3

In both a), b), we can use the identities  $\sin tA = \frac{1}{2i}(e^{t(iA)} - e^{t(-iA)})$  and similarly for the cosine.

7. a

- [b)] We start with this one since we can use it all on the other cases. Write out the exponential series:

$$\begin{aligned}
 e^{iA} &= \sum_{n=0}^{\infty} \frac{(iA)^n}{n!} \\
 &= \sum_{k=0}^{\infty} \frac{(iA)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(iA)^{2k+1}}{(2k+1)!} \\
 &= \sum_{k=0}^{\infty} \frac{(-1)^k A^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k A^{2k+1}}{(2k+1)!} \\
 &= \cos A + i \sin A
 \end{aligned}$$

We see that this holds always.

- [c)] By manipulating series, one sees, as in the previous exercise, that

$$\sin A = \frac{1}{2i}(e^{iA} - e^{-iA})$$

Now, we have

$$\sin(A+B) = \frac{1}{2i}(e^{i(A+B)} - e^{-i(A+B)})$$

Suppose  $AB = BA$ . Then the exponential can be separated:

$$\begin{aligned}
 \sin(A+B) &= \frac{1}{2i}(e^{iA}e^{iB} - e^{-iA}e^{-iB}) \\
 &= \frac{e^{iA} - e^{-iA}}{2i} \frac{e^{iB} + e^{-iB}}{2} + \frac{e^{iB} - e^{-iB}}{2i} \frac{e^{iA} + e^{-iA}}{2} \\
 &= \sin A \cos B + \sin B \cos A
 \end{aligned}$$

Now, if the matrices don't necessarily commute, we may find a counterexample. Take for  $A$  and  $B$  something easy to compute, e.g.

$$A = \begin{bmatrix} 2\pi & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

It is easily checked that the matrices don't commute. Then we have that

$$\begin{aligned}
 \sin(A+B) &= \sin A \cos B + \cos A \sin B \\
 &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
 \end{aligned}$$

whereas if  $A$  and  $B$  were taken in reverse order:

$$\begin{aligned}
 \sin(A+B) &= \sin(B+A) = \sin B \cos A + \cos B \sin A \\
 &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
 \end{aligned}$$

Thus we finally see that  $\sin(A+B)$  is well defined for commuting  $A, B$ , but not necessarily if they don't commute.

- [a)] Just as in the previous case, one can show that for  $\cos(A + B) = \cos A \cos B - \sin A \sin B$  if  $A$  and  $B$  commute. Take  $B = -A$ . Then

$$\begin{aligned} I &= \cos 0 = \cos(B - B) = \cos B \cos B - \sin B \sin(-B) \\ &= (\cos B)^2 + (\sin B)^2 \end{aligned}$$

So this identity holds always.

- [d)] Let  $J$  be the Jordan normal form of the matrix  $A$ . Write  $J = D + N$  where  $D$  is the diagonal part. Then

$$\begin{aligned} e^{2\pi i J} &= e^{2\pi i (D+N)} = e^{2\pi i D} e^{2\pi i N} \\ &= e^{2\pi i D} (I + 2\pi i N + (2\pi i)^2 N^2 / 2 + \dots) \end{aligned}$$

Note that if  $N$  is non-zero, then the matrix  $e^J$  will not be diagonal, and so cannot be equal to  $I$ . We conclude that  $N = 0$ .  $e^D$  is easily calculated - every entry is equal to  $e^{2\pi i \lambda_i}$ , where  $\lambda_i$  is the  $i$ -th diagonal entry (the rest are zero). Thus we need to have

$$e^{2\pi i \lambda_i} = 1$$

for all the  $i$ . We conclude that  $\lambda_i$  must be an integer.

Now take a general matrix  $A$  and let  $P$  the similarity matrix such that  $P^{-1}AP = I$ . Assume that

$$e^{2\pi i A} = I$$

Multiply from the left with  $P^{-1}$  and from the right with  $P$  to find that

$$P^{-1} e^{2\pi i P} P = e^{2\pi i J} = I$$

As already shown, this means that  $J$  must be a diagonal matrix with integer entries. We conclude that a necessary condition is that  $A$  be a diagonalizable matrix with integer eigenvalues.

This is sufficient as well, since then

$$\begin{aligned} e^{2\pi i J} &= I \quad / \quad P \cdot * \cdot P^{-1} \\ e^{2\pi i A} &= I \end{aligned}$$

(we performed conjugation by  $P$  on the first equation.)

- [e)] Assume first that  $A(t)$  commutes with its derivative  $\frac{dA}{dt}$ . Then

$$\frac{dA^n(t)}{dt} = nA^{n-1} \frac{dA}{dt}$$

With that in mind, we calculate the derivative of the matrix exponential:

$$\begin{aligned} \frac{de^{A(t)}}{dt} &= \frac{d}{dt} \sum_{n=0}^{\infty} \frac{A^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{d}{dt} \frac{A^n}{n!} \\ &= \sum_{n=1}^{\infty} \frac{A^{n-1}}{(n-1)!} \frac{dA}{dt} \\ &= \left( \sum_{n=0}^{\infty} \frac{A^n}{n!} \right) \frac{dA}{dt} \\ &= e^{A(t)} \frac{dA}{dt} \end{aligned}$$

What if  $A(t)$  and its derivative don't commute? Take

$$A = \begin{bmatrix} 1 & t \\ 0 & 0 \end{bmatrix}$$

Then

$$\frac{dA(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

and

$$A \frac{dA(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \frac{dA(t)}{dt} A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Note that this means that

$$\begin{aligned} \frac{dA^n(t)}{dt} &= \frac{dA(t)}{dt} A^{n-1} + A \frac{dA(t)}{dt} A^{n-2} + \dots \\ &= 0 + 0 + \dots + A^{n-1} \frac{dA(t)}{dt} \\ &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{dA}{dt} \end{aligned}$$

Thus

$$\frac{de^{A(t)}}{dt} = \sum_{n=1}^{\infty} \frac{1}{n!} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & e-1 \\ 0 & 0 \end{bmatrix}$$

On the other hand

$$\begin{aligned} e^{A(t)} \frac{dA}{dt} &= \left( \sum_{n=0}^{\infty} \frac{A^n}{n!} \right) \frac{dA}{dt} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \frac{dA}{dt} \\ &= \begin{bmatrix} 0 & e \\ 0 & 0 \end{bmatrix} \end{aligned}$$

## Result

7 of 7

- b), a) hold always. c) depends on the matrices satisfying special properties e.g. the matrices should commute.
- d) is true for diagonalizable matrices  $A$  with integer eigenvalues.
- e) Is true if the matrix  $A(t)$  commutes with its own derivative.

8. a



Let  $B_k = [b_{ij}^k]$  (here  $k$  is an index, not an exponential) be a sequence of matrices converging to the matrix  $B = [b_{ij}]$ . This means that the entries of  $B_k$  converge to the corresponding entries in  $B$  i.e.

$$b_{ij}^k \xrightarrow{k \rightarrow \infty} b_{ij}$$

Let  $P = [a_{ij}]$  and  $P^{-1} = [c_{ij}]$ . Then

$$\begin{aligned} P^{-1}B_kP &= (P^{-1}B_k)P \\ &= \left[ \sum_{l=1}^n c_{il}b_{lj}^k \right] P \\ &= \left[ \sum_{m=1}^n \left( \sum_{l=1}^n c_{il}b_{lm}^k \right) a_{mj} \right] \end{aligned}$$

Taking the limit as  $k$  tends to infinity, we find that

$$\begin{aligned} \lim_{k \rightarrow \infty} P^{-1}B_kP &= \left[ \lim_{k \rightarrow \infty} \sum_{m=1}^n \left( \sum_{l=1}^n c_{il}b_{lm}^k a_{mj} \right) \right] \\ &= \left[ \sum_{m=1}^n \left( \sum_{l=1}^n c_{il}b_{lm} a_{mj} \right) \right] \\ &= P^{-1}BP \end{aligned}$$

## Result

The exercise is easily solved by writing out the matrix product explicitly.

# Miscellaneous Problem

1. a

We determine the group  $O_n(\mathbb{Z})$  of orthogonal matrices with integer entries. Let  $A = [a_{ij}]$  be such a matrix. We require that

$$\begin{aligned} AA^t &= I \\ [a_{ij}][a_{ji}] &= [\delta_{ij}] \end{aligned}$$

which amounts to the requirement that

$$\sum_{i=1}^n a_{ik}a_{jk} = \delta_{ij}$$

Taking  $i = j$ , we find that

$$\sum_{i=1}^n a_{ik}^2 = 1$$

meaning that the row vectors have norm 1. Since the  $a_{ik}$  are integers, we conclude that one of them must be  $\pm 1$  and **all** the other zeroes.

Since the same holds for  $A^tA = I$ , we conclude likewise that there is only one non-zero element equal to  $\pm 1$  in each column. Thus we see that  $A \in O_n(\mathbb{Z})$  is necessarily a matrix with  $\pm 1$  exactly once in each row and each column.

Now we show that each such matrix indeed satisfies  $A^t A = I$ .

Let  $A$  be a matrix with  $\pm 1$  exactly once in each row and column. Then if  $i = j$

$$\begin{aligned}\sum_{i=1}^n a_{ik}^2 &= 0^2 + \dots + (\pm 1)^2 + \dots + 0^2 \\ &= 1\end{aligned}$$

and if  $i \neq j$ , then

$$\begin{aligned}\sum_{i=1}^n a_{ik} a_{jk} &= 0 \cdot 0 + \dots + 0 \cdot (\pm 1) + \dots + (\pm 1) \cdot 0 + \dots + 0 \cdot 0 \\ &= 0\end{aligned}$$

Thus we have that

$$\sum_{i=1}^n a_{ik} a_{jk} = \delta_{ij}$$

i.e.  $AA^t = I$ . This shows that  $O_n(\mathbb{Z})$  is made up of "signed" permutation matrices.

## Result

3 of 3

From  $AA^t = I$  we see that  $\sum_{i=1}^n a_{ik}^2 = 1$  and so there can be only one non-zero entry (namely  $\pm 1$ ) per row.  
Since  $AA^t = I$ , the same holds for columns.

It is easy to see that the necessary condition is also sufficient. Thus  $O_n(\mathbb{Z})$  is made up of "signed" permutation matrices.

## 2. a

Let  $J$  be a matrix in Jordan form and suppose it has Jordan blocks:

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

Suppose that the characteristic polynomial of  $J$  is

$$p(t) = (t - \lambda_1)^{k_1} \dots (t - \lambda_m)^{k_m}$$

Then plugging in  $J$ , we get

$$(J - \lambda_1 I)^{k_1} \dots (J - \lambda_m I)^{k_m}$$

Now,  $J - \lambda_i I$  will have a zero diagonal block since the corresponding Jordan block in  $J - \lambda_i I$  will be nilpotent with exponent  $< k_i$ .

Thus, the upper left corner of  $J - \lambda_1 I$  will be zero. Then a successive block will be zero in  $J - \lambda_2 I$  and so on. Thus we have a product of block-diagonal matrices where there is always a zero block in the product. Thus

$$p(J) = 0$$

Now, let  $A$  be an arbitrary complex matrix and  $J$  be its Jordan normal form.  $A$  and  $J$  have the same characteristic polynomial. Then

$$\begin{aligned}
 p(A) &= p(PJP^{-1}) \\
 &= \sum_{i=0}^n a_i (PJP^{-1})^i \\
 &= P \left( \sum_{i=0}^n a_i J^i \right) P^{-1} \\
 &= Pp(J)P^{-1} \\
 &= P0P^{-1} \\
 &= 0
 \end{aligned}$$

## Result

3 of 3

The theorem is first proven for matrices in Jordan normal form for which it is easy to verify. Then it follows for arbitrary matrices by noting that  $p(A) = Pp(J)P^{-1}$ .

### 3. a

Let  $A$  be an  $n \times n$  complex matrix. Let  $J$  be its Jordan normal form. Suppose  $J = P^{-1}AP$ . Then

$$J^k = (P^{-1}AP)^k = P^{-1}A^kP$$

showing that  $J^k$  is the Jordan normal form of  $A^k$ . Considering that  $J$  is a triangular matrix,  $J^k$  has on the diagonal the  $k$ -th powers of the diagonal entries of  $J$ .

We conclude  $\lambda$  is an eigenvalue of  $A$  (and  $J$ ) if and only if  $\lambda^k$  is an eigenvalue of  $A^k$  (and  $J^k$ )

Assume there are  $m$  non-zero eigenvalues. Now, consider the traces of  $A^k$ . Since all these are zero, we arrive at the (infinite) sequence of equalities:

$$\begin{aligned}
 \lambda_1 + \lambda_2 + \dots + \lambda_m &= 0 \\
 \lambda_1^2 + \lambda_2^2 + \dots + \lambda_m^2 &= 0 \\
 \lambda_1^3 + \lambda_2^3 + \dots + \lambda_m^3 &= 0 \\
 &\vdots
 \end{aligned}$$

(with  $\lambda_i \neq 0$ ). The powers appearing above may remind us of the Vandermonde determinant. The determinant

$$\begin{aligned}
 \begin{vmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_m \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^m & \lambda_2^m & \dots & \lambda_m^m \end{vmatrix} &= \lambda_1 \lambda_2 \dots \lambda_m \begin{vmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_m \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \dots & \lambda_m^{m-1} \end{vmatrix} \\
 &= \lambda_1 \lambda_2 \dots \lambda_m \prod_{i \neq j} (\lambda_i - \lambda_j)
 \end{aligned}$$

will be the determinant of the matrix of the system

$$\begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_m \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^m & \lambda_2^m & \dots & \lambda_m^m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

We have chosen the  $\lambda_i$  to be non-zero. We see from the above that the determinant is non-zero, thus  $AX = 0$  should have the unique solution  $X = (0, 0, \dots, 0)$ , the zero vector.

But notice that  $X = (1, 1, \dots, 1)$  also solves the above equation. We have arrived at a contradiction. It must be that there are no non-zero eigenvalues.

Putting  $A$  in Jordan form, it is immediate that  $A = PJP^{-1}$  is nilpotent.

## Result

3 of 3

Note that  $A^k$  has as eigenvalues exactly  $\lambda^k$  where  $\lambda$  is an eigenvalue of  $A$ .

The sequence of equalities naturally gives rise to a Vandermonde determinant. Assuming there are non-zero eigenvalues, we construct a system of equations which has two solutions where it should in fact be unique. This is a contradiction, and proves that all eigenvalues are zero. Hence  $A$  is nilpotent.

## 4. a

Let  $A$  be a complex  $n \times n$  matrix all of whose eigenvalues have absolute value less than 1. Consider the powers of  $A$ . We can use the Jordan normal form, to find that

$$\begin{aligned} A^k &= (PJP^{-1})^k \\ &= PJ^kP^{-1} \\ &= P(D + N)^kP^{-1} \end{aligned}$$

where  $D$  is the diagonal and  $N$  the nilpotent part of the matrix  $J$ . Now,  $N$  is nilpotent of order at most  $n$  (since it is an  $n \times n$  matrix).  $D$  and  $N$  commute, so upon expanding by the binomial formula, we get the sum:

$$\begin{aligned} A^k &= P \left( \sum_{i=0}^k \binom{k}{i} D^{k-i} N^i \right) P^{-1} \\ &= P \left( D^k + kD^{k-1}N + \dots + \binom{k}{n} D^{k-n} N^n + \dots \right) P^{-1} \end{aligned}$$

Note that for large enough  $k$ , the sum ends as soon as we arrive at  $N^n = 0$ , thus the number of summands doesn't depend on  $k$ . Finally, take the limit as  $k$  tends to infinity.

$$\lim_{k \rightarrow \infty} A^k = P \left( \sum_{i=0}^{n-1} \lim_{k \rightarrow \infty} \binom{k}{i} D^{k-i} N^i \right) P^{-1}$$

The binomial coefficient is a polynomial in  $k$  whereas the entries of  $D^k$  are exponentials with base  $< 1$ . Thus the limit results in the zero matrix and we conclude that

$$\lim_{k \rightarrow \infty} A^k = 0$$

Now, consider the partial sums  $1 + A + \dots + A^k$ . Multiplying this with  $I - A$  we find that

$$(I - A)(1 + A + \dots + A^k) = I - A^{k+1}$$

Taking the limit as  $k \rightarrow \infty$ , we find that

$$(I - A)(1 + A + \dots + A^k + \dots) = I$$

Since the inverse of a matrix is unique, we conclude that  $(I - A)^{-1} = 1 + A + \dots + A^k + \dots$ .

## Result

3 of 3

Since the eigenvalues are less than 1, one can prove that  $A^k$  must converge to the zero matrix. Then just note that

$$(I - A)(1 + A + \dots + A^k) = I - A^{k+1} \text{ and take the limit.}$$

## 5. a

Write the Fibonacci recursion as

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} f_{n-2} \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} f_{n-1} \\ f_n \end{bmatrix}$$

- [a] Denote the matrix on the left with  $A$ . It is easily seen (and proven by induction) that

$$A^n \begin{bmatrix} f_0 \\ f_1 \end{bmatrix} = \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix}$$

Now  $f_0 = 0, f_1 = 1$  and so if we can calculate  $A^n$  explicitly we will have derived an explicit formula for  $f_n$  as well. This is done by diagonalizing the matrix. The eigenvalues are:

$$t^2 - t - 1 = 0$$

$$t = \frac{1 \pm \sqrt{5}}{2}$$

leading to the eigenvectors

$$\begin{bmatrix} \frac{-1+\sqrt{5}}{2} \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{-1-\sqrt{5}}{2} \\ 1 \end{bmatrix}$$

We could now calculate  $A^n$ , but we can also just express  $(0, 1)$  in terms of the eigenvectors. We find that

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = a \begin{bmatrix} \frac{-1+\sqrt{5}}{2} \\ 1 \end{bmatrix} + b \begin{bmatrix} \frac{-1-\sqrt{5}}{2} \\ 1 \end{bmatrix}$$

with  $a = -\frac{1+\sqrt{5}}{2\sqrt{5}}$  and  $b = \frac{\sqrt{5}-1}{2\sqrt{5}}$ . Finally, with this:

$$\begin{aligned} \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix} &= A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= A^n \left( -\frac{1+\sqrt{5}}{2\sqrt{5}} \begin{bmatrix} \frac{-1+\sqrt{5}}{2} \\ 1 \end{bmatrix} + \frac{\sqrt{5}-1}{2\sqrt{5}} \begin{bmatrix} \frac{-1-\sqrt{5}}{2} \\ 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} \frac{1}{\sqrt{5}} \left( \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{1-\sqrt{5}}{2} \right)^n \right) \\ \dots \end{bmatrix} \end{aligned}$$

(the second row of the matrix is the same with  $n + 1$  instead). This proves the desired formula (which is incorrectly stated in the textbook).

- [b]) Note that the relation  $a_n = \frac{1}{2} (a_{n-1} + a_{n-2})$  can be written as

$$\begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} a_{n-2} \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix}$$

Call the matrix on the left  $B$ . Then we have that

$$B^n \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$$

The eigenvalues of  $B$  are  $\lambda = 1, -\frac{1}{2}$  with respective eigenvectors

$$(1, 1), (-2, 1)$$

Then writing  $(a_0, a_1)$  in the eigenvector basis, we find that

$$\begin{aligned} \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} &= A^n \begin{bmatrix} a_1 \\ a_0 \end{bmatrix} \\ &= \frac{2a_1 + a_0}{3} A^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{a_1 - a_0}{3} A^n \begin{bmatrix} -2 \\ 1 \end{bmatrix} \end{aligned}$$

Taking the limit as  $n \rightarrow \infty$ , we find that

$$\lim_{n \rightarrow \infty} \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \frac{2a_1 + a_0}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

i.e.

$$\lim_{n \rightarrow \infty} a_n = \frac{2a_1 + a_0}{3}$$

## Result

3 of 3

In both a) and b), we find the eigenvalues and eigenvectors of corresponding matrices. Then write  $(a_1, a_0)$  in the eigenvector basis.

6. a



### Image of $A$

: Let  $A(u, v) = u - v$ . Considering this as a "matrix", one defines the special multiplication of  $A$  and a one-dimensional function  $f(u)$  with

$$A \cdot f = \int_0^1 A(u, v) f(v) dv$$

In this way  $A$  maps from the set of continuous functions (on  $[0, 1]$ ) again to the set of continuous functions. For the above  $A$ , we get that

$$\begin{aligned} (u - v) \cdot f(v) &= \int_0^1 (u - v) f(v) dv \\ &= u \int_0^1 f(v) dv - \int_0^1 v f(v) dv \\ &= au + b \end{aligned}$$

a linear function (note that the integrals evaluate simply to some constants). Thus we see that the image of  $A$  is contained in the set of all linear functions. In fact it is equal to the whole set, for let  $f(u) = Au + B$ . Then

$$(u - v) \cdot (Au + B) = u \left( \frac{A}{2} + B \right) - \left( \frac{A}{3} + \frac{B}{2} \right)$$

Equate this with an arbitrary  $au + b$ . It is sufficient (and necessary) that the coefficients be the same, that is

$$\begin{aligned} \frac{A}{2} + B &= a \\ \frac{A}{3} + \frac{B}{2} &= -b \end{aligned}$$

Note that the determinant of the system is  $\neq 0$ , thus there is a unique solution for any  $a, b$ . This shows that one can always choose  $A, B$  such that  $(u - v) \cdot f(u) = au + b$ .

We conclude that the image of  $A(u, v) = u - v$  is the set of linear functions  $au + b$  (defined on  $[0, 1]$ ) where  $a, b$  are arbitrary.

**Eigenvalues:** We have seen that  $A(u, v)$  maps into linear functions. Thus if there is a function  $f(v)$  such that

$$A(u, v) \cdot f(u) = \lambda f(u)$$

it must certainly be a linear function. Let  $f(u) = au + b$  and suppose it was an eigenvector. Then it must hold that

$$\begin{aligned} A(u, v) \cdot f(u) &= u \left( \frac{a}{2} + b \right) - \left( \frac{a}{3} + \frac{b}{2} \right) \\ &= \lambda f(u) = \lambda(au + b) \end{aligned}$$

Equating coefficients, we arrive at the system of equations

$$\begin{aligned} \frac{a}{2} + b &= \lambda a \\ \frac{a}{3} + \frac{b}{2} &= -\lambda b \end{aligned}$$

It is seen that if  $a = 0$ , then necessarily  $b = 0$  and vice-versa giving the zero function (but this is not a true eigenvector).

Now, assume that  $a, b$  are not equal to zero. Then we may divide out  $a, b$  to get

$$\begin{aligned} \frac{1}{2} + \frac{b}{a} &= \lambda \\ \frac{a}{3b} + \frac{1}{2} &= -\lambda \end{aligned}$$

Adding the two together, we arrive at the necessary condition:

$$1 + \frac{b}{a} + \frac{a}{3b} = 0$$

which is in fact a quadratic. Solving this, we find that

$$a = \frac{-3 \pm \sqrt{-3}}{2} b$$

Plugging this into the second equation of the above system of equations, we find that

$$\begin{aligned} \frac{\frac{-3 \pm \sqrt{-3}}{2} b}{3} + \frac{b}{2} &= -\lambda b \\ \lambda &= \pm \frac{i}{2\sqrt{3}} \end{aligned}$$

Thus the possible eigenvalues are the two values above, and they are indeed achieved (e.g. by taking  $b = 1$  and  $a = \frac{-3 \pm \sqrt{-3}}{2}$  as indicated above).

**Kernel:** Finally, consider the kernel for  $A(u, v) = u - v$ . These are those functions  $f(u)$  such that

$$A(u, v) \cdot f(u) = u \int_0^1 f(v) dv - \int_0^1 v f(v) dv \equiv 0$$

Again, equating coefficients, we find that the kernel of  $A$  is made up of functions satisfying the two conditions:

$$\begin{aligned} \int_0^1 f(v) dv &= 0 \\ \int_0^1 v f(v) dv &= 0 \end{aligned}$$

#### Image of $A$

: Let  $A(u, v) = u^2 + v^2$ . For this  $A$ , we get that

$$\begin{aligned} (u^2 + v^2) \cdot f(v) &= \int_0^1 (u^2 + v^2) f(v) dv \\ &= u^2 \int_0^1 f(v) dv + \int_0^1 v^2 f(v) dv \\ &= au^2 + b \end{aligned}$$

a quadratic function (note that the integrals evaluate simply to some constants). Thus we see that the image of  $A$  is contained in the set of quadratic functions without linear term. In fact it is equal to the whole set, for let  $f(u) = Au^2 + B$ . Then

$$(u^2 + v^2) \cdot (Au^2 + B) = u^2 \left( \frac{A}{3} + B \right) - \left( \frac{A}{5} + \frac{B}{3} \right)$$

Equate this with an arbitrary  $au^2 + b$ . It is sufficient (and necessary) that the coefficients be the same, that is

$$\begin{aligned} \frac{A}{3} + B &= a \\ \frac{A}{5} + \frac{B}{3} &= b \end{aligned}$$

Note that the determinant of the system is  $\neq 0$ , thus there is a unique solution for any  $a, b$ . This shows that one can always choose  $A, B$  such that  $(u^2 + v^2) \cdot f(u) = Au + B$ .

We conclude that the image of  $A(u, v) = u^2 + v^2$  is the set of linear functions  $au^2 + b$  (defined on  $[0, 1]$ ) where  $a, b$  are arbitrary.

**Eigenvalues:** We have seen that  $A(u, v)$  maps into quadratic functions without a linear term. Thus if there is a function  $f(v)$  such that

$$A(u, v) \cdot f(u) = \lambda f(u)$$

it must certainly be such a function. Let  $f(u) = au + b$  and suppose it was an eigenvector. Then it must hold that

$$\begin{aligned} A(u, v) \cdot f(u) &= u \left( \frac{a}{3} + b \right) + \left( \frac{a}{5} + \frac{b}{3} \right) \\ &= \lambda f(u) = \lambda(au + b) \end{aligned}$$

Equating coefficients, we arrive at the system of equations

$$\begin{aligned} \frac{a}{3} + b &= \lambda a \\ \frac{a}{5} + \frac{b}{3} &= \lambda b \end{aligned}$$

It is seen that if  $a = 0$ , then necessarily  $b = 0$  and vice-versa giving the zero function (but this is not a true eigenvector).

Now, assume that  $a, b$  are not equal to zero. Then we may divide out  $a, b$  to get

$$\begin{aligned} \frac{1}{3} + \frac{b}{a} &= \lambda \\ \frac{a}{5b} + \frac{1}{3} &= \lambda \end{aligned}$$

Subtracting the two together, we arrive at the necessary condition:

$$\frac{b}{a} - \frac{a}{5b} = 0$$

which is in fact an easy quadratic. Solving this, we find that

$$a = \pm\sqrt{5}b$$

Plugging this into the second equation of the above system of equations, we find that

$$\begin{aligned} \frac{\pm\sqrt{5}b}{5} + \frac{b}{3} &= \lambda b \\ \lambda &= \frac{1}{3} \pm \frac{1}{\sqrt{5}} \end{aligned}$$

Thus the possible eigenvalues are the two values above, and they are indeed achieved (e.g. by taking  $b = 1$  and  $a = \pm\sqrt{5}$  as indicated above).

**Kernel:** Finally, consider the kernel for  $A(u, v) = u^2 + v^2$ . These are those functions  $f(u)$  such that

$$A(u, v) \cdot f(u) = u^2 \int_0^1 f(v) dv + \int_0^1 v^2 f(v) dv \equiv 0$$

Again, equating coefficients, we find that the kernel of  $A$  is made up of functions satisfying the two conditions:

$$\begin{aligned} \int_0^1 f(v) dv &= 0 \\ \int_0^1 v^2 f(v) dv &= 0 \end{aligned}$$

## Result

7 of 7

For  $A(u, v) = u - v$  it is seen that  $A \cdot f$  results in a linear function. Thus the image is the set linear functions. The eigenvalues are then some specific linear functions. The kernel is found by equating coefficients to zero.

For  $A(u, v) = u^2 + v^2$  similar applies, with the image of  $A$  now being quadratic without a linear term.

## 7. a

Let  $A$  be a  $2 \times 2$  complex matrix with distinct eigenvalues, and let  $X$  be an indeterminate  $2 \times 2$  matrix. Let  $J$  be the Jordan normal form. Then there is a  $P$  such that  $A = PJP^{-1}$ . Plug that into the equation:

$$\begin{aligned} X^2 &= A = PJP^{-1} \quad / \quad P^{-1} \cdot \cdot P \\ P^{-1}X^2P &= J \\ P^{-1}XPP^{-1}XP &= J \\ (P^{-1}XP)^2 &= J \end{aligned}$$

Denote  $P^{-1}XP = Y$ . We have reduced the problem to the same problem for a matrix in Jordan normal form. Since  $X = PYP^{-1}$ , the number of solutions  $Y$  will be the same as that of  $X$ . Let

$$Y = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

. There are two possibilities.

First, suppose that  $A$  has two **distinct** eigenvalues  $\lambda_1 \neq \lambda_2$ . Then we arrive at the matrix equation

$$\begin{bmatrix} a^2 + bc & b(a + d) \\ c(a + d) & d^2 + bc \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Equating the corresponding matrix elements, we arrive at the system:

$$\begin{aligned} a^2 + bc &= \lambda_1 \\ b(a + d) &= 0 \\ c(a + d) &= 0 \\ d^2 + bc &= \lambda_2 \end{aligned}$$

Distinguish two cases:

- If  $a + d = 0$ , this means that  $a = -d$  and so  $a^2 = d^2$ . But then we would have

$$\lambda_1 = a^2 + bc = d^2 + bc = \lambda_2$$

contrary to the assumption that the eigenvalues are distinct. We conclude that this case is impossible.

- Thus it must be that  $a + d \neq 0$ . This immediately implies  $b = c = 0$  and leaves the two equations

$$\begin{aligned} a^2 &= \lambda_1 \\ d^2 &= \lambda_2 \end{aligned}$$

which if solvable have two solutions each. (Note that in no case  $a + d = 0$ .) We conclude that if  $J$ , i.e.  $A$ , has **distinct positive** eigenvalues, then there are 4 solutions to the equation  $X^2 = A$ .

If one of the eigenvalues was 0 and the other positive, then there would be only 2 solutions.

Now, suppose that  $A$  has only one eigenvalue  $\lambda$ . There are two possible Jordan forms. First, consider the matrix equation

$$\begin{bmatrix} a^2 + bc & b(a + d) \\ c(a + d) & d^2 + bc \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

Equating the corresponding matrix elements, we arrive at the system:

$$\begin{aligned} a^2 + bc &= \lambda \\ b(a + d) &= 0 \\ c(a + d) &= 0 \\ d^2 + bc &= \lambda \end{aligned}$$

Take  $a + d = 0$ . Then the second and third equations are satisfied and the fourth becomes equivalent to the first. Thus we are left with

$$a^2 + bc = \lambda$$

Choosing  $a$  and  $c \neq 0$  arbitrarily, we take  $b = \frac{\lambda - a^2}{c}$ . Thus we see that in this case there is an **infinite** number of solutions.



Lastly, suppose that  $A$  has only one eigenvalue  $\lambda$  and  $J$  is of the other possible form. Consider the matrix equation

$$\begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{bmatrix} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

Equating the corresponding matrix elements, we arrive at the system:

$$\begin{aligned} a^2 + bc &= \lambda \\ b(a+d) &= 1 \\ c(a+d) &= 0 \\ d^2 + bc &= \lambda \end{aligned}$$

Now  $a+d \neq 0$  because of the second equation and so the third equation can hold if and only if  $c = 0$ . One immediately gets

$$\begin{aligned} a^2 &= \lambda \\ d^2 &= \lambda \end{aligned}$$

Since  $a+d \neq 0$ , we conclude that  $a, d$  must have the same sign, and so there are the two possibilities:

$$a = d = \pm\sqrt{\lambda}$$

(note that it must be that  $\lambda > 0$  since  $a+d \neq 0$ ).  $b$  is explicitly calculated now as  $b = \frac{1}{a+d}$ , and so we conclude that there are two possibilities.

## Result

5 of 5

Is easy to reduce the equation  $X^2 = A$  to an equation  $Y^2 = J$  with the same number of solutions, where  $J$  is the Jordan normal form.

If  $A$  has distinct eigenvalues, they must be non-negative and there are 4 solutions if they are both positive, and 2 if one is zero.

If  $A$  has one eigenvalue and is diagonalizable, then are an infinity of solutions in any case.

If  $A$  has one eigenvalue but is not diagonalizable, then the eigenvalue must be positive and there are two solutions to the equation.

## 8. a

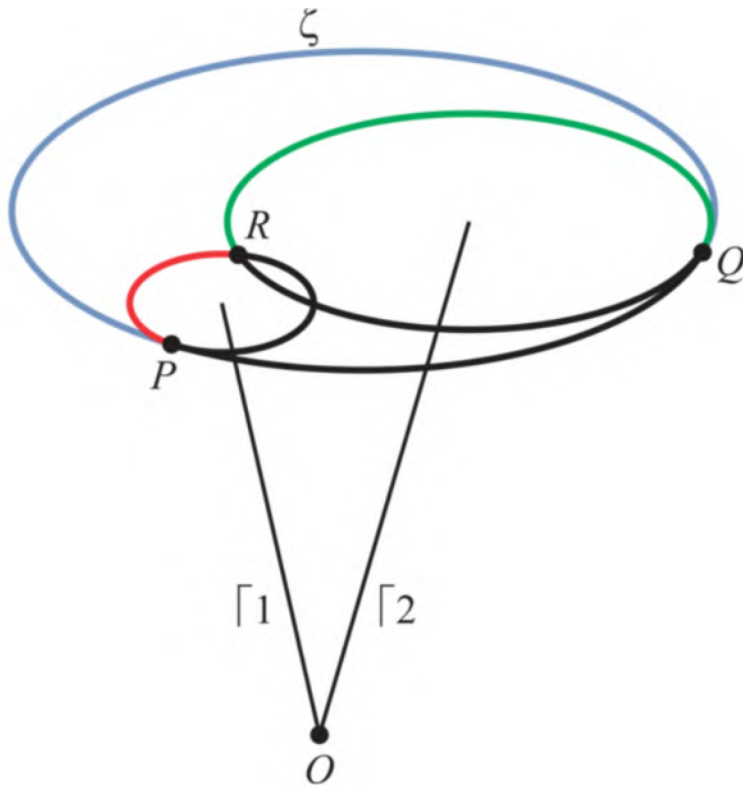
To determine the axis of rotation for the composition of two three dimensional rotations by geometrically;

[Comment](#)

Step 2 of 3 ^

Draw the following diagram as shown below:





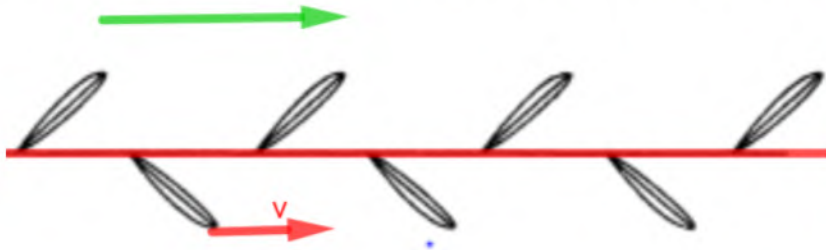
- 1: Draw circles of arbitrary size around the axes of rotation.
- 2: Take the point  $P$  on the first circle that, when rotated by the red angle  $\theta_1$ . Clockwise about  $\Gamma_1$ , is mapped onto the closer of the two points of intersection of the two circles. This point of intersection is now rotated by the green arc  $\theta_2$  about  $\Gamma_2$  to get the final image point  $Q$ .
- 3: Construct a circle  $\zeta$  tangent to the first circle at  $P$  and the second circle at  $Q$ .
- 4: The line joining the center of  $\zeta$  to the origin is the axis of rotation, and the angle of the composed rotation is the one subtended by the blue arc.

## Chapter 6

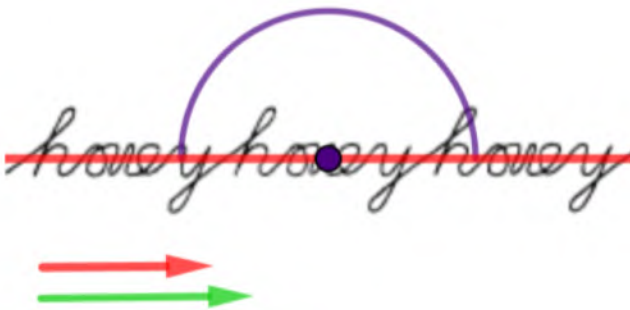
### Section 1

1. a

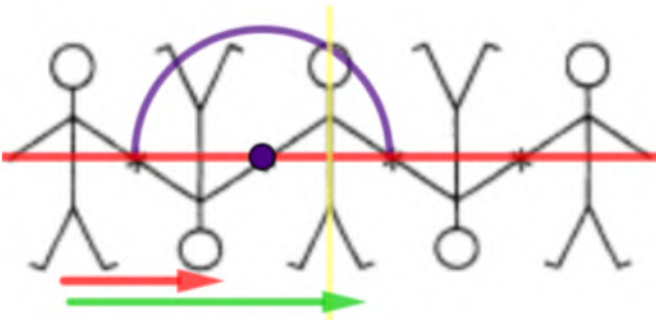
Besides the **glide** symmetry noted in the textbook (red line and vector), there is also a **translational** symmetry (green vector). No matter what axis is chosen, there is no rotational or bilateral symmetry.



There is a **translational** symmetry by translating one *h* to the next *h*. Similarly, first reflecting around the red axis and translating, one finds a **glide** symmetry. Finally, there is also a **purple** rotational symmetry which can be achieved by rotating around the center of the letter *n* by  $180^\circ$ . There is no bilateral symmetry.



There is a **translational** symmetry by translating one man to the next. Also, one can reflect them by the red line and translate by the red vector for a **glide** symmetry. Rotating around the purple point for a  $180^\circ$  will yield a **rotational** symmetry. Finally, there is a **bilateral** symmetry given by reflecting the image by the yellow axis in the middle of a man.



## Result

For the first picture, we find a glide and translational symmetry.

For the second there is a translational, glide and rotational symmetry.

For the last picture, all symmetries are present.

## Section 3

### 1. a

We verify the rules for composition of plane isometries given in the textbook. Recall that for orthogonal operators, it was shown that  $\varphi t_a = t_{\varphi(a)}\varphi$ . Since rotation by an angle  $\theta$  around the origin is an orthogonal (linear) operator, we have that

$$\rho_\theta t_v = t_{\rho_\theta(v)}\rho_\theta$$

Note that the reflection  $r$  around the  $e_1$ -axis is also an orthogonal matrix (e.g. look at its matrix). Therefore the same logic as above shows that

$$r t_v = t_{r(v)}r$$

Finally, for rotations and the reflection  $r$ , we can use their matrices to find that

$$\begin{aligned} r\rho_\theta &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & -\cos \theta \end{bmatrix} \\ \rho_{-\theta}r &= \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & -\cos \theta \end{bmatrix} \end{aligned}$$

showing that indeed  $r\rho_\theta = \rho_{-\theta}r$ .

That  $t_v t_w = t_{v+w}$  was shown before. That the composition of two rotations behave as given is obvious, but it is also not hard to prove by matrix multiplication:

$$\begin{aligned} \rho_\theta \rho_\eta &= \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \eta & -\sin \eta \\ \sin \eta & \cos \eta \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta \cos \eta - \sin \theta \sin \eta & -(\sin \theta \cos \eta + \sin \eta \cos \theta) \\ \sin \theta \cos \eta + \sin \eta \cos \theta & \cos \theta \cos \eta - \sin \theta \sin \eta \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta + \eta) & -\sin(\theta + \eta) \\ \sin(\theta + \eta) & \cos(\theta + \eta) \end{bmatrix} \\ &= \rho_{\theta+\eta} \end{aligned}$$

and it is completely obvious that  $r^2 = 1$  ( $r$  is a diagonal matrix with entries  $\pm 1$ ). Thus we have proven all the rules given in the textbook.

## Result

2 of 2

Recall that for an orthogonal operator  $\phi t_a = t_{\phi(a)}\phi$ . Now, note that rotation and reflection are orthogonal operators. That  $r\rho_\theta = \rho_{-\theta}r$  and  $\rho_\theta\rho_\eta = \rho_{\theta+\eta}$  is shown by matrix multiplication. Also, it had already been shown that  $t_v t_w = t_{v+w}$  and it is obvious that  $r^2 = 1$ .

## 2. a

Suppose  $m$  is an orientation-reversing isometry. Then it is either a reflection or a glide reflection (reflection + translation by a vector  $v$  parallel to the line of reflection). This means that

$$m = t_v r$$

where  $v$  may possibly be the zero vector (then it is just a reflection). Now follows a simple calculation:

$$\begin{aligned} m^2 &= t_v r t_v r \\ &= t_v t_{r(v)} r r \\ &= t_v t_v \cdot 1 \\ &= t_{2v} \end{aligned}$$

Thus we see that  $m^2$  is a translation.

## Result

2 of 2

From the rule that  $r t_v = t_{r(v)} r$ , it follows quickly that  $m^2$  must be a translation.

## 3. a

Let  $A$  be a linear operator on  $\mathbb{R}^2$  that is a reflection about the line

$$ax + by + c = 0$$

The normal vector for this line is given by  $(a, b)$ . Perpendicular to this vector is e.g. the vector  $(-b, a)$ . Geometrically,  $(a, b)$  is reversed by  $A$  and  $(-b, a)$  is preserved, so that

$$\begin{aligned} A \begin{bmatrix} a \\ b \end{bmatrix} &= -1 \begin{bmatrix} a \\ b \end{bmatrix} \\ A \begin{bmatrix} -b \\ a \end{bmatrix} &= 1 \begin{bmatrix} -b \\ a \end{bmatrix} \end{aligned}$$

Thus we see that  $A$  has the eigenvalues  $1, -1$ , with eigenvectors which are orthogonal.

Now, suppose that  $A$  was a linear operator on  $\mathbb{R}^2$  and that its eigenvalues were 1 and  $-1$ . Eigenvectors corresponding to different eigenvalues must be orthogonal. Suppose  $(a, b)$  was an eigenvector for  $-1$ . Then  $(b, -a)$  must be an eigenvector for 1. We may as well take  $a^2 + b^2 = 1$ . Consider now the change of basis matrix

$$P = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

This is a matrix for a reflection. Since reflection is self-inverse, we have  $P^{-1} = P$ . Thus

$$\begin{aligned} A &= P^{-1}DP \\ &= \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \\ &= \begin{bmatrix} -a^2 + b^2 & -2ab \\ -2ab & -b^2 + a^2 \end{bmatrix} = \begin{bmatrix} c & s \\ s & -c \end{bmatrix} \end{aligned}$$

which has the form of a reflection. Finally, indeed it is a reflection since the determinant is

$$-(-a^2 + b^2)^2 - 4a^2b^2 = -a^4 - 2a^2b^2 - b^4 = -(a^2 + b^2)^2 = -1$$

## Result

3 of 3

If  $A$  reflects about the line  $ax + by + c = 0$ , then it is easily seen that  $(a, b)$  is eigenvector  $-1$  and  $(-b, a)$  an eigenvector for 1.

If  $A$  is a linear operator with eigenvalues 1,  $-1$ , we can take  $(a, b)$  and  $(b, -a)$  to be the orthogonal eigenvectors (also take  $a^2 + b^2 = 1$ ). It is not hard to write down the change of base matrix and find the general form of  $A$ . We see that it is a reflection matrix.

## 4. a

Recall that the group of isometries is generated by the translation  $t_a$ , rotations  $\rho_\theta$  and the reflection  $r$ . Any glide reflection can thus be written as  $t_a\rho_\theta r$ . We check that the conjugates with the generators satisfy the requirements.

- Let  $t_v$  be an arbitrary translation. Suppose  $\vec{i}$  is a vector in the gliding direction and  $\vec{j}$  is a vector in the perpendicular direction. This is a basis and we can write  $v = (v_1, v_2)$  and  $a = (a_1, a_2)$  in this basis. Note that only the first component corresponds to the gliding vector, which is  $(a_1, 0)$ . Then

$$\begin{aligned} t_v^{-1}t_a\rho_\theta r t_v &= t_{-v}t_a t_{\rho_\theta r(v)} r \\ &= t_{-v+a+\rho_\theta r(v)} r \end{aligned}$$

Now,  $\rho_\theta r(v) = (v_1, -v_2)$  since  $\rho_\theta r$  is a reflection around the axis of the  $\vec{i}$  vector. Thus

$$t_v^{-1}t_a\rho_\theta r t_v = t_{(a_1, 0)} t_{(0, a_2 - 2v_2)} \rho_\theta r$$

Note that the gliding vector still has the same length  $a_1$  along the  $\vec{i}$  direction.

- Now, take a rotation  $\rho_\eta$ . Then

$$\begin{aligned} \rho_\eta^{-1}t_a\rho_\theta r \rho_\eta &= t_{r\rho_\eta^{-1}(a)} \rho_{-\eta+\theta} \rho_{-\eta} r \\ &= t_{r\rho_\eta^{-1}(a)} \rho_{-2\eta+\theta} r \end{aligned}$$

Geometrically, one can see that the composition  $\rho_{-2\eta+\theta} r$  will be the reflection around the axis of the reflection  $r$  rotated by

$$\frac{1}{2}(-2\eta + \theta) = -\eta + \theta/2$$



or equivalently, this axis is rotated  $-\eta$  degrees from the axis of  $\rho_\theta r$ .

It follows that a gliding vector for  $\rho_{-2\eta+\theta}r$  must be rotated by  $-\eta$  from a gliding vector for  $\rho_\theta r$ . This is exactly the case for  $\rho_{-\eta}(a)$ .

Thus the gliding vector  $\rho_{-\eta}(a)$  is the same vector just rotated by  $-\eta$ , and therefore of the same length.

- Finally, taking  $r$ , it is easy to find that

$$\begin{aligned} r^{-1}t_a\rho_\theta r r &= r t_a \rho_\theta \\ &= t_{r(a)}\rho_{-\theta}r \end{aligned}$$

Note that the axis of the glide reflection  $\rho_{-\theta}r$  is on the opposite side of the axis for  $\rho_{-\theta}r$ . Therefore, the vector  $a$  must be reflected over that axis to become a gliding vector for  $a$ . This is exactly the case for  $r(a)$ . We conclude that  $r(a)$  has the same length as  $a$ .

Now, taking a general element of the group of isometries  $M$ , one can write it out as product of translations, rotations and  $r$ . We prove the general statement by mathematical induction. The base case is the one above where  $m$  is one of the generating elements.

Suppose we have proven the statement when  $m$  is a product of at most  $n - 1$  generating elements. Now, consider a product  $p_1 p_2 \dots p_n$ , where the  $p_i$  are translations, rotations or  $r$ . Then

$$(p_1 p_2 \dots p_n)^{-1} t_a \rho_\theta r (p_1 p_2 \dots p_n) = p_n^{-1} \underbrace{(p_1 \dots p_{n-1})^{-1} t_a \rho_\theta r (p_1 \dots p_{n-1})}_{\text{is a glide reflection by induction}} p_n$$

Denoting this glide reflection with  $g$ , we also have that  $p_n^{-1} g p_n$  is a glide reflection.

By the principle of mathematical induction, we have proven the statement for a product of arbitrary length. Thus we have proven the statement for all elements of  $M$ .

## Result

3 of 3

One can write glide reflections as  $t_a \rho_\theta$ . Note that the vector  $a$  need not be the glide vector - in fact it can be written as  $a = (a_1, a_2)$  where the first component is in the direction of the axis of  $\rho_\theta r$ .

With this in mind, one checks that  $m^{-1} t_a \rho_\theta r m$  is a glide reflection for any generator of the group of isometries  $M$ .

5. a



Let  $x = x_1 + ix_2$  and  $a = a_1 + ia_2$ . The complex number  $x$  will correspond to the vector/point  $(x_1, x_2)$ . The elementary symmetries from the textbook now become:

- translation  $t_a$  by a vector  $a$

$$\begin{aligned} t_a(x) &= x + a = (x_1 + ix_2) + (a_1 + ia_2) \\ &= x_1 + a_1 + i(x_2 + a_2) \end{aligned}$$

- rotation  $\rho_\theta$  by an angle  $\theta$  around the origin. Recall that multiplication of complex numbers multiplies their absolute values and adds arguments. Thus multiplying with  $e^{i\theta} = \cos \theta + i \sin \theta$  simply rotates the number by  $\theta$ . We find that

$$\begin{aligned} \rho_\theta(x) &= (\cos \theta + i \sin \theta)(x_1 + ix_2) \\ &= x_1 \cos \theta - x_2 \sin \theta + i(\cos \theta x_2 + x_1 \sin \theta) \end{aligned}$$

- reflection around  $e_1$

$$r(x) = x_1 - ix_2 = \overline{x}$$

where the overline denotes complex conjugation.

## Result

2 of 2

In terms of complex numbers, translation becomes addition, rotation multiplication by a complex number of absolute value 1, and reflection is just complex conjugation.

## 6. a

- [a] Let  $s$  be the rotation of the plane with angle  $\pi/2$  about the point  $(1, 1)^t$ . This transformation can be realized by first shifting the center of rotation into the origin. After that one rotates and just translates back by  $(1, 1)$ . Thus the transformation in question is equal to

$$\begin{aligned} s &= t_{(1,1)} \rho_{\pi/2} t_{(-1,-1)} \\ &= t_{(1,1)} t_{\rho_{\pi/2}(-1,-1)} \rho_{\pi/2} \\ &= t_{(1,1)} t_{(1,-1)} \rho_{\pi/2} \\ &= t_{(2,0)} \rho_{\pi/2} \end{aligned}$$

- [b)] Let  $s$  denote reflection of the plane about the vertical axis  $x = 1$ . By inspection, we find that

$$s(x, y) = (2 - x, y)$$

This can be written as

$$\begin{aligned} s(x, y) &= t_{(2,0)}(-x, y) \\ &= t_{(2,0)}\rho_{\pi}(x, -y) \\ &= t_{(2,0)}\rho_{\pi}r(x, y) \\ s &= t_{(2,0)}\rho_{\pi}r \end{aligned}$$

and so we have found an expression for  $s$ . To find the conjugating element  $g$ , we recall that  $t_a^{-1} = t_{-a}$  and  $\rho_{\theta}^{-1} = \rho_{-\theta}$ . Thus  $g$  and  $g^{-1}$  shouldn't look too different. We try the following trick:

$$s = t_{(1,0)}t_{(1,0)}\rho_{\pi/2}\rho_{\pi/2}r$$

Now, we switch over one of the translations and one of the rotations to the other side. We have

$$\begin{aligned} s &= t_{(1,0)}t_{(1,0)}\rho_{\pi}r \\ &= t_{(1,0)}\rho_{\pi}t_{(-1,0)}r \\ &= t_{(1,0)}\rho_{\pi/2}\rho_{\pi/2}rt_{(-1,0)} \\ &= t_{(1,0)}\rho_{\pi/2}r\rho_{-\pi/2}t_{(-1,0)} \end{aligned}$$

We see that by taking  $g = t_{(1,0)}\rho_{\pi/2}$ , we arrive at an element such that  $s = grg^{-1}$ .

## Result

3 of 3

- It is easy to see that the same transformation is realized by first translating everything to the origin, rotating and translating back. The rest is calculation.
- Note that  $s(x, y) = (2 - x, y) = t_{(2,0)}\rho_{\pi}r(x, y)$ . This can be rewritten to find the required conjugating element  $g$  to be  $t_{(1,0)}\rho_{\pi/2}$ .

# Section 4

## 1. a

Recall the defining relations for the dihedral group  $D_n$ :

$$\begin{aligned} x^n &= 1 \\ y^2 &= 1 \\ xy &= yx^{-1} \end{aligned}$$

Using these repeatedly we bring the given expression into the form  $x^i y^j$ . This goes like

$$\begin{aligned} x^2(yx^{-1})y^{-1}x^3y^3 &= x^2(xy)y^{-1}x^3y^2y \\ &= x^3 \cdot 1 \cdot x^3 \cdot 1 \cdot y \\ &= x^6y \end{aligned}$$

## Result

2 of 2

Use the definition of the abstract dihedral group  $D_n$ . The important rule is that  $xy = yx^{-1}$ .

## 2. a

- [a] We list all the subgroups of  $D_4 = \{1, x, x^2, x^3, y, xy, x^2y, x^3y\}$ . We do this by sorting them by size.
- The subgroup of size 1 is of course  $\{1\}$ . It is normal.
- The subgroups of size 2 must be cyclic i.e. generated by elements of order 2. These are the reflections of which there are 4 and rotation by  $180^\circ$ :

$$\{1, x^2\}, \\ \{1, y\}, \{1, xy\}, \{1, x^2y\}, \{1, x^3y\}$$

Let  $z$  is the non-trivial element in the above subgroups. The subgroups will be normal if  $gzg^{-1}$  is again in the subgroup. But reflections don't commute with rotations.

Thus the only possible normal subgroup  $H = \{1, x^2\}$ . It is easily checked that this is so. (It is enough to check that  $x^{-1}Hx = H$  and  $y^{-1}Hy = H$  since  $x, y$  are the generators.)

- Groups of order 4 are abelian, either cyclic or generated by two elements of order 2. Suppose  $x$  or  $x^3$  is in the group. Then all its powers are also in it, and so we get the subgroup

$$\{1, x, x^2, x^3\}$$

Suppose  $y$  is in the subgroup.  $xy, x^3y$  together with  $y$  generate the whole group, thus a subgroup of order 4 with  $s$  must be

$$\{1, s, r^2, sr^2\}$$

Finally, if  $r, r^3, s$  are not in the subgroup, we are left with

$$\{1, r^2, sr, sr^3\}$$

which is indeed a subgroup.

Subgroups of index 2 are always normal subgroups, and so the same applies to the above.

- Finally, the whole group  $D_4$  is normal in itself.

- [b] Consider the dihedral group  $D_{15}$ . Suppose  $N$  was a normal subgroup. Recall that every reflection is of the form  $x^n y$  for some  $n = 0, 1, \dots, 14$ . Suppose a reflection was in  $N$ . Then for  $N$  to be normal, we would need

$$(x^k)^{-1} x^n y x^k = x^{n-2k} y$$

for every  $k$ . Since  $\gcd(2, 15) = 1$ , the expression  $x^{n-2k} y$  can be any reflection  $y, xy, \dots, x^{14}y$ . Thus a normal subgroup with reflections would contain **all** the reflections. Since it must contain additionally the identity, we would have a subgroup with  $\geq 16$  elements - showing that  $N = G$ .

Thus a **proper** normal subgroup  $N$  must not contain reflections. We conclude that this must be the group of rotations  $N = \{1, x, \dots, x^{14}\}$ .

- [c] We list the subgroups of  $D_6$  not containing  $x^3$ . Obviously, such a subgroup cannot contain  $x, x^3, x^5$ . Also note that

$$x^n y x^m y = x^{n-m}$$

and therefore such a subgroup must not contain two reflections  $x^n y, x^m y$  with  $n, m$  of different parity (since then the result will be one of  $x, x^3, x^5$ ).

We see that the subgroups may contain  $x^2, x^4$  and reflections  $x^n y, x^m y$  with the same parity. This gives the subgroups:

$$\begin{aligned} &\{1, y\}, \{1, xy\}, \{1, x^2 y\}, \{1, x^3 y\}, \{1, x^4 y\}, \{1, x^5 y\} \\ &\{1, x^2, x^4\}, \\ &\{1, x^2, x^4, y, yx^2, yx^4\}, \{1, x^2, x^4, yx, yx^3, yx^5\} \end{aligned}$$

## Result

4 of 4

- Subgroups of order 4 (i.e. of index 2) are always normal. There is only one normal subgroup of order 2, that is  $\{1, x^2\}$ .
- We show that a normal subgroup of  $N$  containing reflections must be the whole group. There a proper normal subgroup must be the group of rotations.
- Note that such a subgroup cannot have  $x, x^3, x^5$  or  $x^n, x^m y$  with  $n, m$  of the same parity. With this in mind, it's not hard to write out all the groups.

## 3. a

### Step 1

1 of 4

- [a] Let  $H = \{1, x^5\}$ . This is a subgroup of  $D_{10}$ . We write down the left cosets. This is done by simply calculating  $zH$ , for  $z \in D_{10}$  until we exhaust all possibilities. The cosets are:

$$\begin{aligned} H &= \{1, x^5\} \\ xH &= \{x, x^6\} \\ x^2H &= \{x^2, x^7\} \\ x^3H &= \{x^3, x^8\} \\ x^4H &= \{x^4, x^9\} \\ yH &= \{y, yx^5\} \\ yxH &= \{yx, yx^6\} \\ yx^2H &= \{yx^2, yx^7\} \\ yx^3H &= \{yx^3, yx^8\} \\ yx^4H &= \{yx^4, yx^9\} \end{aligned}$$

Every element of  $D_{10}$  is listed in one of the above cosets, therefore this is all of them.

- [b)] We prove that  $H$  is normal. We must show that  $g^{-1}Hg = H$  for arbitrary  $g$ . If  $g = x^n$ , then

$$g^{-1}Hg = x^{-n}Hx^n = \{1, x^5\} = H$$

for any  $n$ . If  $g = yx^n$ , then

$$g^{-1}Hg = (yx^n)H(yx^n) = x^{-n}y\{1, x^5\}yx^n = \{1, x^{-5}\} = \{1, x^5\} = H$$

since  $x^{10} = 1$ . This shows that  $g^{-1}Hg = H$  for all  $g \in D_{10}$  and so  $H$  is normal in  $D_{10}$ .

(Note: it was also sufficient to just check  $g = x, y$  since they generate  $D_{10}$ . The process is the same.)

### Step 3

3 of 4

- [c)] We compare the elements of order 2. In any dihedral group, reflections are always of order 2. The only rotation  $x^n$  of order 2 is  $x^5$  in  $D_{10}$ . Thus there are 11 elements of order 2 in  $D_{10}$ .

Similarly, in  $D_5$  there are 5 reflections. There is no rotation by  $180^\circ$ , and so there are in total 5 elements of order 2. For each of these reflections, we have two elements in  $D_5 \times H$  (since  $H$  has two elements). This gives a total of at most 10 elements of order 2.

We conclude that the two groups are not isomorphic.

### Result

4 of 4

a) Calculate  $zH$  for  $z \in D_{10}$  until exhausting all elements.

b) It is checked directly that  $g^{-1}Hg = H$ . One can also consider only  $g = x, y$ .

- c) Compare the number of elements of order 2. There are 10 reflections plus 1 rotation by  $180^\circ$  in  $D_{10}$ . But in  $D_5 \times H$  there are at most 10 such elements.

## Section 5

1. a



Let  $\ell_1$  and  $\ell_2$  be lines through the origin in  $\mathbb{R}^2$  that intersect in an angle  $\pi/n$ , and let  $r_i$  be the reflection about  $\ell_i$ .

Recall that  $D_n$  is generated by a reflection and rotation. Since we already have two reflections, we need only show that we can generate a rotation  $\rho_{2\pi/n}$ .

Consider the effect of reflecting a point over  $\ell_2$ . The same effect can be achieved by rotating the point by  $-\pi/n$ , reflecting it over  $\ell_1$  and then rotating it back by  $\pi/n$ . Thus

$$r_2 = \rho_{-\pi/n} r_1 \rho_{\pi/n}$$

Using the rules for manipulating such expressions, we find that

$$\begin{aligned} r_2 &= r_1 \rho_{\pi/n} \rho_{\pi/n} \\ r_1^{-1} r_2 &= \rho_{2\pi/n} \end{aligned}$$

Thus  $r_1, r_2$  generate a rotation by  $2\pi/n$ .

$r_1$  and  $\rho_{2\pi/n}$  generate a dihedral group,  $D_n$ , as required.

## Result

2 of 2

Notice that the other reflection can be written as  $r_2 = \rho_{-\pi/n} r_1 \rho_{\pi/n}$ . From this we can express  $\rho_{2\pi/n}$ . This rotation and either of the reflections  $r_1, r_2$  generate  $D_n$ .

## 2. a

Consider a discrete group of isometries whose translation group  $L$  has the form  $\mathbb{Z}a$  with  $a \neq 0$ . Let  $H$  be the crystallographic restriction. Recall that the crystallographic restriction contains rotations of order 1, 2, 3, 4, 6.

Since  $H$  operates on the lattice  $L$  as well, which is one-dimensional, it must be that  $H$  contains at most rotations of order 1 or 2. This gives the possibilities  $C_1, C_2 = D_1$  and  $D_2$ . All happen as can be demonstrated:

$$\begin{array}{c} \circ \square \triangle \circ \square \triangle \circ \square \triangle \\ \square \triangle \square \triangle \square \triangle \\ \square X \square X \square X \end{array}$$

Note that all of the above patterns have one-dimensional discrete translational symmetry.

The first one has a trivial crystallographic restriction  $H = C_1$ , since no (vertical) axis can be chosen for reflection to be a symmetry and it is evident that rotation by  $180^\circ$  will rotate the triangle.

The second one can be reflected around a vertical axis, but cannot be rotated - thus it has  $H = C_2$ .

The third one can be reflected around a vertical and horizontal axis. Note that a composition of such reflections is exactly a rotation by  $180^\circ$ . Thus here  $H = D_2$ .

## Result

2 of 2

Since  $\mathbb{Z}a$  is one-dimensional, the crystallographic restriction  $H$  can contain at most the rotation by  $180^\circ$  and no other ones. We show that  $H$  can be the either  $C_1, C_2$  or  $D_2$ .

## 3. a



Let  $L = \mathbb{Z}a + \mathbb{Z}b$  be a lattice in  $\mathbb{R}^2$  generated by the vectors  $a, b$ . A sublattice  $L'$  is a lattice generated by some vectors in  $L$ .

By writing  $xa + yb = (x, y)$ , we may without loss of generality assume that the lattice is  $L = \mathbb{Z}^2$ .

## Step 2

2 of 4

Suppose  $L'$  contains the vector  $(x, 0)$  and let  $x$  be the minimal positive integer with this property. Consider the quotient

$$L/L'$$

For any element  $v + L'$  of the quotient group, we can consider

$$(\alpha, 0) + v + L' \in L/L', \quad 0 \leq \alpha < x$$

For  $0 \leq \alpha < x$ , we will get different cosets (else we would have  $(\alpha, 0) \in L'$  contradicting the fact that  $x$  is the smallest such number).

We conclude that  $L/L'$  can be divided into  $x$  equal partitions. This shows that  $x$  divides the index  $|L/L'|$  of  $L'$  in  $L$ .

Let  $n = |L/L'|$ . We conclude that it suffices to check all sublattice  $L'$  with vectors of the form  $(d, 0)$ , where  $d$  is a divisor of  $n$ . We find that \* if  $(1, 0) \in L'$  then any vector  $(a, b) \in L'$  can be transformed to  $(0, b)$  and vice-versa. It is easy to see that  $L'$  will have index in  $L$  equal to  $b$ . Thus  $L'$  can be expressed as

$$L' = \mathbb{Z}(1, 0) + \mathbb{Z}(0, 3)$$

- if  $(3, 0) \in L'$ , then consider a vector  $(a, b) \in L'$ . By adding  $(3, 0)$ , we can reduce that vector to either  $(0, b)$ ,  $(1, b)$  and  $(2, b)$ . Consider the lattices

$$L' = \mathbb{Z}(3, 0) + \mathbb{Z}(0, 1)$$

$$L' = \mathbb{Z}(3, 0) + \mathbb{Z}(1, 1)$$

$$L' = \mathbb{Z}(3, 0) + \mathbb{Z}(2, 1)$$

All of these have index 3. If  $b > 1$ , then it will be immediately seen that the lattice would have greater index.

We see that these are all the lattices, giving 4 in total.

## Result

4 of 4

Introducing coordinates, we may assume that the lattice is  $\mathbb{Z}^2$ .

First we show that if  $(x, 0) \in L'$  with  $x$  being the minimal positive integer with this property, then  $x$  divides the index of  $L'$  in  $L$ .

This shows that it suffices to choose  $(d, 0) \in L'$  with  $d$  a divisor of  $n$ . Then it's not hard to choose another vector so we have a basis for  $L'$ .

4. a

Let  $(a, b)$  be a lattice basis of a lattice  $L$  in  $\mathbb{R}^2$ . Let  $(a', b')$  be some other basis. Since they are bases, we must have that there are integers such that

$$\begin{aligned} a' &= \alpha a + \beta b \\ b' &= \gamma a + \delta b \end{aligned}$$

which can also be written in the matrix form

$$(a', b') = (a, b)P, \quad P = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

Note that these  $a', b', a, b$  are vectors, and so in fact  $(a', b')$  and  $(a, b)$  are matrices.

Similarly, there is a matrix  $Q$  with integers entries such that  $(a, b) = (a', b')Q$ . This gives

$$\begin{aligned} (a', b') &= (a, b)P = (a', b')QP \\ (QP - I)(a', b') &= 0 \end{aligned}$$

Now, the matrix  $(a', b')$  is composed of two linearly independent columns, and so has determinant  $\neq 0$ .

Multiplying by its inverse, we find that

$$\begin{aligned} QP - I &= 0 \\ QP &= I \\ \det(Q) \det(P) &= 1 \end{aligned}$$

Since  $P, Q$  were matrices with integers entries,  $\det(P)$  must be an integer and since it divides 1, we conclude that  $\det(P) = \pm 1$ .

## Result

2 of 2

There are integer matrices  $P, Q$  such that  $(a', b') = (a, b)P$  and  $(a, b) = (a', b')Q$ . Then  $(QP - I)(a', b') = 0$  and so  $QP = I$ . Taking determinants, it follows that  $\det(P) = \pm 1$ .

## 5. a

Consider the frieze pattern



Consider first the translational symmetry group  $L$  of this pattern. Obviously, we can translate it left and right. But there cannot be two linearly independent vectors in  $L$  since the pattern is one-dimensional.

We conclude that the translation group  $L = \mathbb{Z}a$  for some vector  $a$  (equal to the difference between two consecutive triangles).

Now, consider the crystallographic restriction. Since the pattern is one dimensional, we know that the only possibilities are rotation by  $180^\circ$  and reflection. While it is possible to reflect the pattern horizontally, it is impossible to rotate it by  $180^\circ$ . We conclude that the crystallographic restriction is  $C_2$ .

This shows that the whole group of symmetries is made of elements of the form  $t_a r$ , where  $t_a$  are translation by multiples of  $a$  and  $r$  is horizontal reflection. Thus

$$G = C_2 \times C_\infty$$

## Result

2 of 2

There is obviously one-dimensional translational symmetry. It is easily that the pattern only has horizontal reflection symmetry. Thus every symmetry is of the form  $t_a r$ , and so the group is  $C_2 \times C_\infty$ .

## 6. a

Let  $G$  be the group of symmetries of the frieze pattern in the textbook. Recall that the point group records the angles of rotations (and slopes of glide lines of reflection). There is an obvious vertical reflection.

Now, put a point in the middle of one of the signs in the pattern. One can do a  $180^\circ$  degree rotation around that to get the same shape again. Since the translational group is one-dimensional, there can be no other rotations. This implies that the point group is

$$\overline{G} = D_2$$

Note that the horizontal reflection is not evident since this is the point group - in fact there is a glide symmetry made of horizontal reflection and translation. Since the point group  $\overline{G}$  is made by quotienting out translations, this is no surprise.

### Step 2

2 of 3

Recall that the point group is derived as the image of the homomorphism

$$\pi : G \rightarrow \overline{G}$$

with translations as kernel. By the first homomorphism theorem, we have that

$$G / \ker \pi = G / T = \overline{G} = D_2$$

showing that the index of the translation group  $T$  in  $G$  is equal to  $|D_2| = 4$ .

### Result

3 of 3

Note that there is a vertical reflection, rotation by  $180^\circ$  degrees and a glide reflection. There can be more symmetries in the point group and so  $\overline{G} = D_2$ .

The index of the translation group is equal to the size of  $\overline{G}$ .

## 7. a

Groups of isometries of a line;

The group  $I(\mathbb{R})$  of isometries of the line  $\mathbb{R}$  is an interesting and (for such an apparently easy case) complicated group.

[Comment](#)

### Step 2 of 3

Let  $N$  denote the group of isometries of a line  $\mathbb{R}$

To classify: the discrete subgroups of  $N$ , identifying those that differ in the choice of origin and unit length on the line

Since the group;

$$O(1) = \{\pm 1\}$$

Then divide the elements of the group into two subsets: those with orthogonal part  $+1$  and those with the orthogonal part  $-1$ .

The former are just translations of the form  $x \mapsto a + x$  while the latter are maps of the form  $x \mapsto a - x$ .

There is always a bijection from the additive group that will lead to the translations.

These later maps have the effect of "reversing the direction" on the line.

Then;

$$x = Ta$$

If there is a reflection on another point;

$$\begin{aligned} R(x) \circ Ry(a) &= Rx(Ry(a)) \\ &= Rx(2y - x) \end{aligned}$$

And;

$$\begin{aligned} Tx \circ Rx(a) &= Tx(Rx(a)) \\ &= Tx(2y - a) \\ &= 2x + y - a \\ a &= R\left(y + \left(\frac{x}{2}\right)\right)(a) \end{aligned}$$

Further;

$$\begin{aligned} R(x) \circ T(y(a)) &= Ra(x + b) \\ &= 2y + x \end{aligned}$$

Reflection in the point  $b$  is the map  $x \mapsto 2b - x$

**Therefore, the required discrete subgroup is  $x = Ta$**

## 8. a

Isometry is defined as the transformation that is invariant with respect to distance. The distance between any two points in the pre-image should be the same like it is between the images two products.

a.

Consider the isometric

$$g \in E, g \in E$$

Mapped as;

$$\begin{aligned} [x, y, z] &\mapsto [y - x, y, -z + 12][x, y, z] \\ &\mapsto [y - x, y, -z + 12] \end{aligned}$$

Order  $gag$  is 2

$$1\{1, g\}\{1, g\} = \text{finite subgroup}$$

Let  $K$  have finite order  $n$ . Consider an orbit  $x$

$$f(x), f(f(x)), \dots, f(n-1)(x), f(x), f(f(x)), \dots, f(n-1)(x)$$

Average is fixed

$$\begin{aligned} y &= x + f(x) + f(f(x)) + \dots + f(n-1)(x)ny \\ &= x + f(x) + f(f(x)) + \dots + f(n-1)(x)n \end{aligned}$$

b.

To make: A careful case analysis for the proof

A frieze group is one of the mathematical concepts which is used in the classification of the design in a two-dimensional space in a space of one direction,

This situation is very common in architecture and decorative art.

Symmetry can occur.

Frieze groups for a two-dimensional line groups, on a one direction, these are related with complex wallpaper groups, repetitive of two directions, and crystallographic groups, repetitive of three directions.

The  $G$  is the subgroup of the Euclidean group with a linear transformation  $T$

Translational subgroup is

$$T \supset GT \supset G$$

These are the isometrics of pure translations

9. a

Let  $G$  be a discrete subgroup of  $M$  whose translation group is not trivial. Consider the possible isometries.

Translations and glide reflections don't have fixed points. Rotations have one. Reflections have a line as the set of fixed points.

Since the group is discrete, there are countably many translation vectors  $t_a$  and the point group is finite. This means that there are countably many different elements  $t_a \rho_\theta r$ . Since each of them has at most a line as a fixed point, we conclude that the set of fixed points  $F$  is a **countable union of lines and points**.

## Step 2

2 of 3

Now, let  $S$  be the set of all slopes of the lines in  $F$ . There are countably many such slopes. But there are uncountably many possible slopes in  $[0, 2\pi)$ . Thus taking  $\alpha \in [0, 2\pi) \setminus S$ , we show that the line  $y = \alpha x$  cannot be in  $F$ .

Suppose it was. Since its slope is different from all lines in  $F$ , it intersects every line once. Since there are countably many lines and points in  $F$ , this shows that there are only countably many points on  $y = \alpha x$  in  $F$ . Since a line has uncountably many points, we conclude that there is a point on the line  $y = \alpha x$ , which is not the set of fixed points for  $G$ .

Thus there is a point in  $\mathbb{R}^2$  which is not fixed by any element of  $G$ .

## Result

3 of 3

Since  $G$  is discrete, its cardinality is at most countable. It is seen that individual isometries have at most a line as a fixed point set (reflections do). Thus the fixed point set is a union of countably many lines (and points). But this cannot be the whole plane.

10. a



Let  $f$  and  $g$  be rotations about distinct points,  $f$  by  $\theta$  and  $g$  by an angle of  $\phi$ . Consider the point group  $\overline{G}$  which is arrived at by mapping the translations to the identity:

$$\pi : G \rightarrow \overline{G}$$

In the point group  $\overline{G}$ ,  $\pi(f) = \overline{f}$  and  $\pi(g) = \overline{g}$  are simply rotations (and the point of reference is irrelevant). Thus imagining they are in the same point, we find that

$$\overline{fgf^{-1}g^{-1}} = \overline{\rho_{\theta}\rho_{\phi}\rho_{-\theta}\rho_{-\phi}} = \overline{1}$$

This means that  $fgf^{-1}g^{-1} \in \ker \pi$ . Thus it is either a translation or the identity.

If it were the identity, then we would have

$$\begin{aligned} fgf^{-1}g^{-1} &= 1 \\ fg &= gf \end{aligned}$$

Let  $F$  be the point about which  $f$  rotates. Then

$$\begin{aligned} gf(F) &= g(f(F)) = g(F) \\ fg(F) &= f(g(F)) \end{aligned}$$

This shows that  $f(g(F)) = g(F)$  and since the only fixed point of a rotation is the point around which we rotate, we would have  $g(F) = F$ . But then  $g$  would rotate around the same point as  $f$ , contrary to assumption that the points are distinct.

We conclude that  $fg \neq gf$  and so  $fgf^{-1}g^{-1}$  is a translation.

## Result

2 of 2

Note that in the point group  $\overline{fgf^{-1}g^{-1}} = \overline{1}$ . Thus  $fgf^{-1}g^{-1}$  is either the identity or a proper translation. Let  $F$  be the point around which  $f$  rotates. Then it's seen that  $fg$  and  $gf$  act differently on  $F$ . Thus  $fgf^{-1}g^{-1} \neq 1$ , and so it must be a translation.

## 11. a

- [a] Let  $\Gamma$  be a subgroup of  $\mathbb{R}^+$ . Suppose that  $G$  is not discrete. Then there is a translation  $t_a$  for which

$$|a| < \epsilon$$

for any choice of  $\epsilon$ . Thus  $\Gamma$  contains elements with arbitrarily small translations. Iterating this translation  $t_a$  on  $0 \in \Gamma$ , we find that

$$t_a^n(0) = na \in \Gamma$$

where  $n \in \mathbb{Z}$ . Let  $b \in \mathbb{R}$  be arbitrary and choose a positive small  $a$ . Then taking  $n = \lfloor \frac{b}{a} \rfloor$ , we have that

$$\begin{aligned} \left\lfloor \frac{b}{a} \right\rfloor a &\leq \frac{b}{a} \cdot a < \left( \left\lfloor \frac{b}{a} \right\rfloor + 1 \right) a \\ 0 &\leq b - na < a \end{aligned}$$

Taking  $a$  to be smaller than  $\epsilon$ , we have that

$$|b - na| < \epsilon$$

where epsilon is arbitrary.

Thus if the group  $\Gamma$  is not discrete, it must be dense in  $\mathbb{R}$ .



- [b)] Suppose a subgroup  $\Gamma$  of  $\mathbb{R}$  contains 1 and  $\sqrt{2}$ . We show that we can get arbitrarily small numbers in  $\Gamma$  and  $\Gamma$  will not be discrete. By part a), it will follow that  $\Gamma$  is dense.

Note that we can get

$$a_1 = \sqrt{2} - 1 = 0.414 \dots \in \Gamma$$

Similarly,

$$a_2 = (\sqrt{2} - 1)^2 = 3 - 2\sqrt{2} = 0.172 \dots \in \Gamma$$

In general, notice that the  $n$ -th power of  $(\sqrt{2} - 1)^n \in \Gamma$  since we can always write out the expression by the binomial theorem:

$$a_n = (\sqrt{2} - 1)^n = \sum_{k=0}^n \binom{n}{k} \sqrt{2}^k (-1)^{n-k} = a + b\sqrt{2} \in \Gamma$$

Obviously,  $a_n$  is a positive sequence converging to 0. Therefore, it gets arbitrarily close to 0, that is, for every  $\epsilon > 0$  there is an  $n$  such that

$$|a_n| < \epsilon$$

This shows that  $\Gamma$  is not discrete, and therefore is dense in  $\mathbb{R}$ .

- [c)] Let  $H$  be a subgroup of the group  $G$  of angles. This means the group

$$G = \mathbb{R}/2\pi\mathbb{Z}$$

i.e. the reals modulo  $2\pi$ . We have the following cases:

- Suppose  $H$  has finitely many angles of the form

$$2\pi \cdot \frac{a}{b}, \quad a, b \in \mathbb{Z}$$

Then this is contained in the cyclic group  $C_n$ , generated by  $\frac{2\pi}{n}$  for some integer  $n$  (take  $n$  to be the least common multiple of all the denominators).

- Suppose  $H$  has infinitely many angles of the above form. If there was a largest possible denominator, then this would not be possible (since there are finitely many fraction with a given denominator). We conclude that the denominators must become arbitrarily large. If  $2\pi \cdot \frac{a}{b} \in H$  with  $\gcd(a, b) = 1$ , then we can also get  $\frac{2\pi}{b}$ . With a large enough denominator, any real number can be approximated by rationals. We conclude that  $H$  is dense in  $G$ .
- Suppose  $H$  contains a number of the form

$$2\pi r$$

where  $r$  is irrational. Consider the multiples of this:

$$r, 2r, 3r, \dots$$

and their fractional parts. Divide the interval  $[0, 1)$  into

$$\left[0, \frac{1}{M}\right), \left[\frac{1}{M}, \frac{2}{M}\right), \dots, \left[\frac{M-1}{M}, 1\right)$$

Consider  $na$  for  $n \in \{1, 2, \dots, M, M+1\}$ . Since there are  $M$  intervals, it must be for at least two  $n_1, n_2$

$$n_1 a, n_2 a$$

must be in the same interval. But then

$$|(n_1 - n_2)a| < \frac{1}{M}$$

Since the choice of  $M$  was arbitrary, we see that we can arbitrarily approximate zero. This shows that  $H$  is not discrete, and therefore  $H$  is dense in  $\mathbb{R}$ .

## Result

4 of 4

a) If  $\Gamma$  is not discrete, then there is an arbitrarily small  $a$  which translates  $\Gamma$ . Since  $0 \in \Gamma$ , we can get any number as

$$t_a^n(0) = na.$$

b) Note that  $(\sqrt{2} - 1)^n \in \Gamma$ . This converges to zero, and so the group is not discrete.

c) If the group of angles contains finitely many angles, then it is cyclic. If there are infinitely many rational angles, or an irrational angle, then the group is dense in  $\mathbb{R}$ .

## 12. a

This theorem is proven by imitating the analogous textbook result for the plane. Let  $L$  be a discrete subgroup of  $\mathbb{R}^+$ .

We show that bounded region of space contains only finitely many points of  $L$ .

Indeed, with  $L$  being discrete, the elements of  $L$  are separated by a distance at least  $\epsilon$ . Take a cube with diagonal  $\epsilon$ . Then there cannot be two elements in that cube. But a bounded region of space can be covered by finitely many cubes, thus there are finitely many points in the bounded region.

If  $L$  is not trivial, then there is some vector  $a$  in it. Consider the circle around the origin of radius  $|a|$ . Since this is a bounded space, there are finitely many points in it. Taking the point closest to the origin yields the vector of minimal length in  $L$ .

If  $B = (u, v, w)$  is a basis for  $\mathbb{R}^3$ , then we can write all vectors  $v \in \mathbb{R}^3$  as

$$v = x + v_0$$

where  $v_0$  is a vector in the fundamental parallelepiped  $\Pi'(B)$  generated by the basis  $B$  and  $x \in L$ .

The rest proceeds similar to the textbook.

•

Suppose all the vectors of  $L$  lie on a line. A line is isomorphic to  $\mathbb{R}$  and so a discrete subgroup must be of the form  $L = \mathbb{Z}a$ , for some vector  $a$ .

- Suppose the vectors lie in a plane but not on a line. Since a plane is isomorphic to  $\mathbb{R}^2$ , the textbook resolves this, since we know that such a discrete subgroup must be  $a\mathbb{Z} + b\mathbb{Z}$ .
- Finally, suppose the vectors don't lie in a plane. Then there are three linearly independent vectors  $a', b', c'$ .

Take two linearly independent vectors  $a', b'$  in  $L$ . They lie in a plane  $P$ . By the second case above,  $P \cap L$  is a lattice of the form  $a\mathbb{Z} + b\mathbb{Z}$ . By changing coordinates, we may assume that  $a, b$  are in fact  $(1, 0, 0), (0, 1, 0) \in L$ .

Next, we replace  $c' = (c'_1, c'_2, c'_3)^t$  by  $-c'$  if necessary, so that  $c'_3$  becomes positive. We look for a vector  $c = (c_1, c_2, c_3)^t$  in  $L$  with  $c_3$  positive, and otherwise as small as possible. A priori, we have infinitely many elements to inspect. However, since  $c'$  is in  $L$ , we only need to inspect the elements  $c$  such that  $0 < b_3 \leq b'_3$ . Moreover, we may add multiples of  $a, b$  to  $c$ , so we may also assume that  $0 \leq c_1, c_2 < 1$ . When this is done,  $c$  will be in a bounded region that contains finitely many elements of  $L$ . We look through this finite set to find the required element  $c$ , and we show that  $B = (a, b, c)$  is a lattice basis for  $L$ .

Let  $\tilde{L} = \mathbb{Z}a + \mathbb{Z}b + \mathbb{Z}c$ . Then  $\tilde{L} \subseteq L$ . We must show that every element of  $L$  is in  $\tilde{L}$ . It is enough to show that the only element

of  $L$  in the region  $\Pi'(B)$  is the zero vector. Let  $d = (d_1, d_2, d_3)$  be a point of  $L$  in that region, so that  $0 < d_1, d_2 < 1$  and  $0 < d_3 < c_3$ . Since  $c_3$  was chosen minimal, it must that  $d_3 = 0$  and the vector  $d$  is in the plane generated by  $a, b$ . But since  $0 \leq d_1, d_2 < 1$ , the only option is that  $d$  is the zero vector.

We conclude that the lattice is of the form  $\mathbb{Z}a + \mathbb{Z}b + \mathbb{Z}c$ .

## Result

3 of 3

The proof is analogous to the textbook proof for  $\mathbb{R}^2$ . Calling on that and finishing the proof analogously, we have that the possible lattices are  $\mathbb{Z}a + \mathbb{Z}b + \mathbb{Z}c$  (with eventually some of  $a, b, c$  being zero).

## Section 6

1. a

(a)

The point group of each and every pattern is considered to be the mirror image of each pattern  
Pattern disclosure or the explanation of the image is as follows

**Group 1**

The first group is transactions only group. Structure or lattice is similar to a parallelogram with translational axis inclined at a specific angle. This will be the exact mirror image no matter how long the pattern goes.

**Group 2**

This is both transactional and rotational they revolve or rotate at a half turn that is 180degree.  
The lattice is an transaction of an angle at an parallelogram. They generally have dot in the figure

**Group 3**

This includes the reflection and translation. They are of two type's reflection which is parallel also known as bilateral symmetry. The other type is in which the lattice is square in shape.

**Group 4:** This is combination of the reflex and transaction this is parallel to one axis and perpendicular to another.

**Group 5:** Reflection and glide reflection have parallel axis and translation.

**Group 6:** Reflection whose axis are perpendicular and rotates into a half turn

**Group7:** Has reflection and glide reflection along with transaction

**Group 8:** Contain glide reflection half turn rotation and transaction.

**Group 9** is reflection and 180degree rotation rhombus lattice

**Group10:** This includes rotation and translation but the rotation is 90 degree.

**Group 11:** Rotation translation and reflection and the rotation century lies reflex axis

**Group12:** contain reflection glide rotation and rotation

**Group 13:** Rotation and translation

**Group14:** similar to 13 but hexagon lattice

**Group 15:** difference from 14 is rotation lies on the axis.

**Group 16** rotation and transactions 60 degree 180degree and half turn Lattice hexagon.

(b)

The patterns that can co-ordinates be chosen so that the group  $G$  operates on the lattice  $L$ .

**Group1:** Structure or lattice is similar to a parallelogram with translational axis inclined at a specific angle. **Group 2:** The lattice is a transaction of an angle at a parallelogram. They generally have dot in the figure.

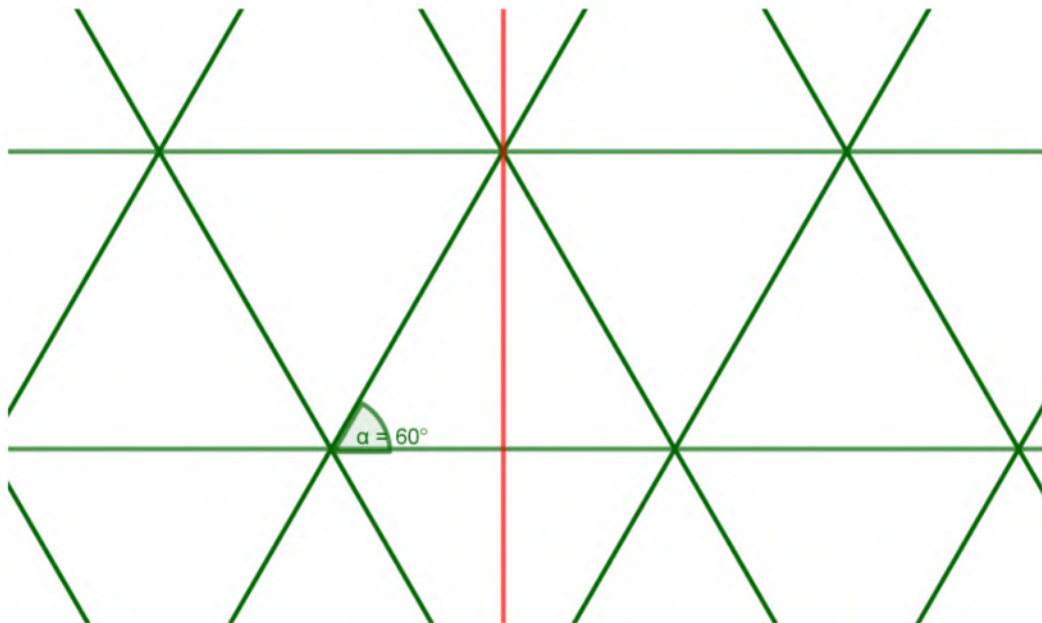
**Group 4** This is combination of the reflex and transaction this is parallel to one axis and perpendicular to another. **Group 5:** Reflection and glide reflection have parallel axis and translation with a rhombus lattice.

**Group 6, 7, 8** has rectangular lattice. **Group 9** is reflection and 180degree rotation rhombus lattice. **Group 11:** Rotation translation and reflection and the rotation century lies reflex axis

**Group14:** similar to 13 but hexagon lattice. **Group 16:** rotation and transactions  $60^\circ, 80^\circ$  and half turn Lattice hexagon.

2. a

Let  $G$  be the group of symmetries of an equilateral triangular lattice  $L$ . Such a lattice will have rotations by  $60^\circ$ . It also has a vertical reflection. Obviously, no other rotations are possible. This means that with rotations and reflections, the point group must be  $D_6$ .



Recall the homomorphism

$$\pi : G \rightarrow \overline{G}$$

which has as kernel the subgroup of translations. By the first isomorphism theorem, the index of this subgroup in  $G$  is

$$|G / \ker \pi| = |D_6| = 12$$

### Result

3 of 3

The symmetries of the triangle lattice are rotations by  $60^\circ$  and reflections. Thus the point is  $D_6$  and the index of the translation group is  $|D_6| = 12$ .

3. a



The top left pattern has rotational symmetry as a square (i.e. rotations by  $90^\circ$ ). No reflections give the same pattern. We conclude that the point group is  $G = C_4$ .

This corresponds to the second pattern in the second row (on the list of 17 patterns).

## Step 2

2 of 5

The top right pattern has rotational symmetry as a square and also has bilateral symmetry (i.e. reflections). This gives the point group  $G = D_4$ .

This corresponds to the fourth pattern in the fourth row (on the list of 17 patterns).

## Step 3

3 of 5

The lower left pattern has hexagons as well as squares. It can either have rotations by  $60^\circ$  or  $90^\circ$ , but not both (else one could also rotate it by  $30^\circ$ , but that is a rotation of order 12.)

It is not possible to find a rotation by  $60^\circ$ . On the other hand, put a point in the center of the square and rotate by  $90^\circ$ . One finds that this preserves symmetry.

This corresponds to the fourth pattern in the fourth row (on the list of 17 patterns).

Horizontal and vertical reflections also preserve the pattern. Thus the point group  $G$  is  $D_4$ .

The lower right pattern has fish with faces. Notice e.g. the eyes. Since they are black down, and white up, there can be no horizontal reflection symmetry and no rotational symmetry. The only symmetry is a vertical glide reflection. Thus the point group consists of one reflection and so  $G = D_2$ .

This corresponds to the second pattern in the first row (on the list of 17 patterns).

## Result

5 of 5

We analyze the patterns by looking for the possible rotational symmetries (rotations by 180, 120, 90, 60 degrees) and reflections. Top left has  $G = C_4$ , top right and lower left has  $G = D_4$  and lower right  $G = C_2$ .

## 4. a

Consider the provided statement to classify plane crystallographic groups with point  $D_1$ .

It is noticed that the discrete subgroups  $k$  in an isometrics of the plane  $k$  with translation lattice  $l$  and it has 2 independent vectors, this too with the point group  $G$  and the dihedral group  $D_1$ . This also contains identity along with reflection along the origin.

[Comment](#)

### Step 2 of 3 ^

From the ten possible groups of point groups  $C_n$  or  $D_n$  where  $n = 1, 2, 3, 4, 6$  out of which  $D_1$  analytics is most complicated. There are 3 different types of group along this point group.

It is assume that,  $G$  -Group of the type that choosing coordinates for the reflection in  $G$  which is alongside to horizontal axis.

A bar over symbols represents elements of the point group  $G$ , the reflection in  $G$  by  $r$ . The lattice  $L$  consists of the vectors  $v$  such that  $TV$  is in  $G$  and Elements of  $G$  map  $L$  to  $L$ . If  $v$  is in  $L$ , then  $RV$  will be also there  $L$  and the shape of the lattice.



### Proposition1

Horizontal and vertical vectors  $a = (a, t, 10)t$  and  $b = (0, b2), c = 12(a + b)$ . One out of two lattices  $L_1$  or  $L_2$ ,

Then  $L_1 = Za + Zb$  are rectangular lattice and  $L_2 = Za + Zc$  is a 'triangular' lattice.

Since  $b = 2c - a$  then  $L_1$  implies to  $L_2$ .  $L_1$  is rectangular because the horizontal and vertical lines along the point divide the pane into a rectangular shape.  $L_2$  is addition of  $L_2$  at the midpoints of rectangles.

5. a

A lattice defined as the algebraic structure whose two elements have a unique supremum or a unique infimum.

[Comment](#)

Step 2 of 4 ^

a.

Consider  $G$  to be a group of isometries of equilateral triangles  $L$ .

Suppose,  $T$  be the group of translations

To prove, that  $G/T = D_6$

If the proof is obtained then the result is done

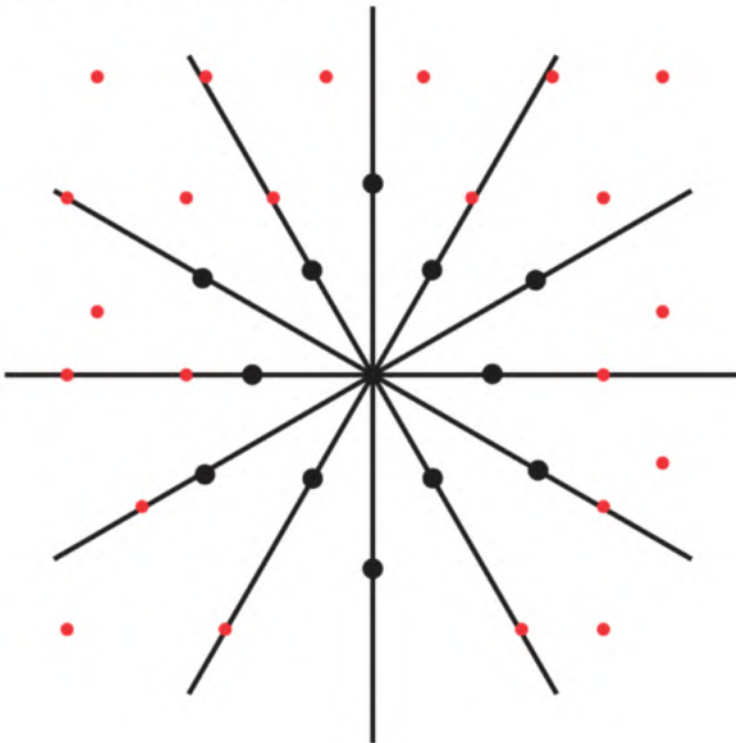
Let if possible suppose;

$$G/T = D_3$$

It is clear that for an equilateral triangle lattice to be formed there should be at least 6 lines that reflect about the horizontal line

But the group  $D_3$  does not contain any reflection about any horizontal line.

Now, consider the below given lattice;



If the lattice is rotated around any single vertex by  $60^\circ$  in such a way that one lattice maps to another lattice point, the composition of 6 such rotation makes the equilateral triangle lattice

Thus, this makes a contradiction for the assumption that;

$$G/T \cong D_3$$

Therefore, the translation group  $L$  is an equilateral triangular lattice if the group of a two-dimensional crystallographic group is  $C_6$  or  $D_6$

b.

Since, there is a small rotation in  $G$  which have finite order say,  $n$

Also, the subgroup of that rotation is isomorphic to  $C_6$  or  $D_6$

This implies that;

$$n = 2$$

So, further by using the theorem of classification of Rosette groups which states that;

A rosette group  $G$  is either a cyclic group  $C_n$  generated by a rotation or is one of the dihedral groups  $D_n$

Rosette groups are mainly composed of the rosette patterns; here the rosette pattern is described as the equilateral triangular lattice

Thus, by using the result of above stated theorem **the required group is Rosette group.**

6. a

A lattice defined as the algebraic structure whose two elements have a unique supremum or a unique infimum.

[Comment](#)

Step 2 of 9 ^

Groups are always defined in a two dimensional plane.

The two features introduced here is the existence of two independent translations corresponding to the two dimensional plane.

The entire pattern is formed by translating the figure into two different directions.

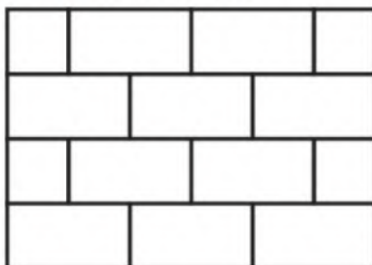
The basic positions obtained by the given figure form the vertices of the infinite lattice.

The lattice of the infinitely many congruent parallelograms as reflected is defined as the 'unit cell'

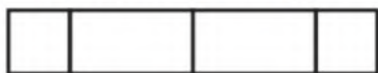
Now, consider any one figure from the given figures;

First separate the figure into two translations.

For this, first consider the figure;



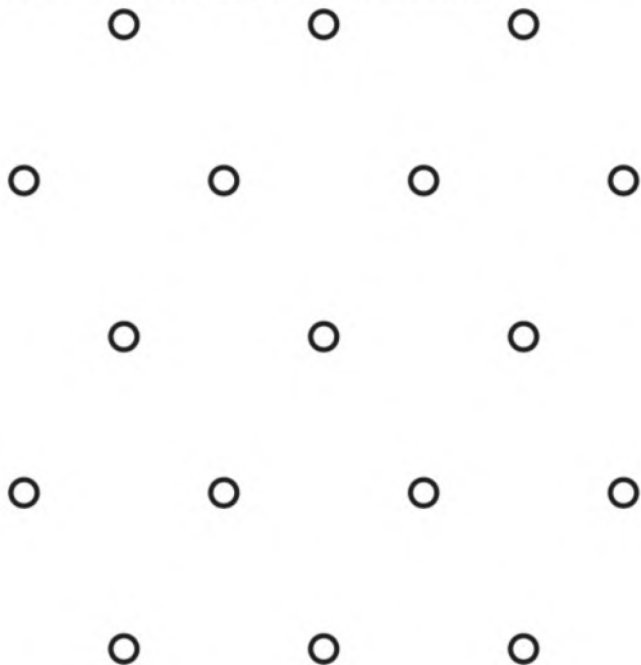
Now, divide this into two translations as done below;



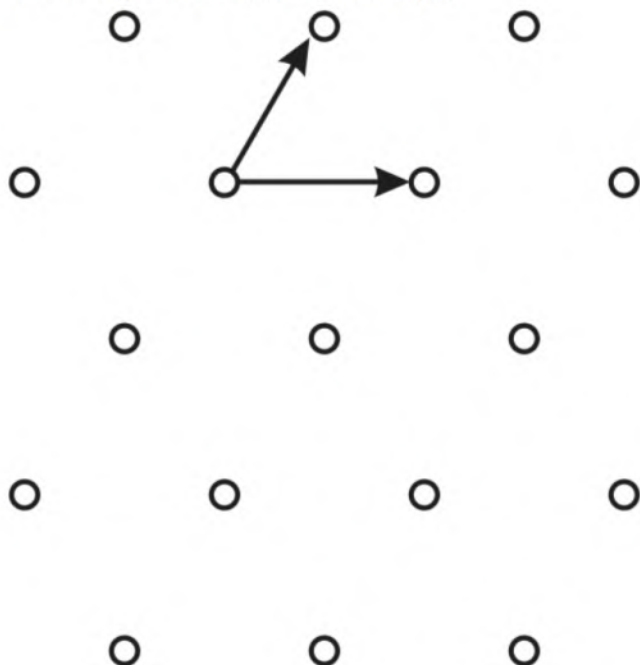
And,



Here, just two colors create entirely new pattern from the basic square lattice. The main lattice of the above is considered as the given diagram of the check board;



For translations, observe the arrows marked on the figure;



Clearly two independent translations can be seen in the plane.

That is the choice of the unit cell is not unique.

This implies that different parallelogram may be taken or even non-parallelograms can be taken.

This contradicts the conditions for a group of symmetry

Similarly, by using the same process for the other given figures it can be obtained that the conditions for a symmetry of group is not satisfied

**Therefore, the given figures exhaust the possibilities.**

## Section 7

### 1. a

Let  $G = D_4$  be the dihedral group of symmetries of the square.

- [a] We calculate the stabilizer of the vertices. This is the subgroup of  $G_s$  of  $G$  such that  $g(v) = v$ . No rotation stabilizes a vertex. But a reflection through that point does. Therefore

$$G_v = \{1, r_v\}$$

This means that the stabilizer group for a vertex is  $C_2$ , where the non-trivial element represents the reflection  $r_v$  about the diagonal through the vertex  $v$ .

Consider now an edge  $e$ . Rotations obviously don't stabilize it. Neither do diagonal reflections. But a reflection through the midpoint of an edge does stabilize it. Call this reflection  $r_e$ . The stabilizer is thus

$$G_e = \{1, r_e\}$$

- [b] Consider  $G$  acting on the set  $\{d_1, d_2\}$  of diagonals. A rotation by  $90^\circ$  degrees obviously changes the diagonals, but a rotation by  $180^\circ$  doesn't. Also reflections around the diagonal itself, or the perpendicular diagonal, leave the diagonal unchanged. Calling the diagonal reflections  $r_{d_1}, r_{d_2}$ , so

$$G_{d_1} = G_{d_2} = \{1, \rho_\pi, r_{d_1}, r_{d_2}\} \cong D_2$$

### Result

3 of 3

The stabilizers of vertices and of edges are isomorphic to  $C_2$ , whereas the diagonals have stabilizer groups isomorphic to  $D_2$ .

### 2. a

The group  $M$  of isometries of the plane operates on the set of lines in the plane. We consider the possible stabilizing elements of lines:

- Any vector on the line (i.e. parallel to the line) translates the line in the same line. Any other vector will translate the line somewhere else.
- Any rotation around a point not on the line moves the line. The only rotation is a rotation around a point on the line by  $180^\circ$ .
- The only reflections are around the line itself and reflections around perpendicular lines.
- There are no non-trivial glide reflections.

Fix a point  $P$  on the line. The translations are  $t_v$  for  $v$  parallel to the line. The rotations are  $t_v \rho_\theta$  where  $\rho_\theta$  is a rotation around  $P$ . Call  $r$  the reflection around the perpendicular line through  $P$ . Then  $t_v r$  is a reflection around another perpendicular line. Thus we have that  $G$  is made of the different elements

$$t_v, \quad t_v \rho_\pi, \quad t_v r, \quad t_v \rho_\pi r$$

for all possible vectors  $v$  parallel to the line. Therefore the stabilizer is

$$G = \{t_v \rho_\pi^i r^j \mid v \text{ parallel to the line, } i = 0, 1, j = 0, 1\} \\ \cong \mathbb{R} \times D_2$$

## Result

2 of 2

Notice that translations with vectors on the line keeps the line fixed. The only other isometries are rotations  $\rho_\pi$  on the line itself and reflections around the line itself and perpendicular lines. One sees that  $G \cong D_2$ .

## 3. a

Let  $U, V$  be two sets of order 3. Suppose the symmetric group  $S_3$  operates on  $U, V$ . Then  $S_3$  naturally acts on  $U \times V$  with the diagonal action

$$g(u, v) = (gu, gv)$$

- [a)] Since the elements of  $U, V$  are irrelevant, we may as well write

$$U = V = \{1, 2, 3\}$$

$S_3$  acts transitively on  $U$ . If all 2-cycles were to act trivially on  $U$ , then  $S_3$  itself would act trivially since it is generated by 2-cycles.

Similarly, if only one 2-cycle, call it  $y$ , was to act non-trivially, then we would have (renaming the elements if necessary)

$$\begin{aligned} y \cdot 1 &= 2 \\ y \cdot 2 &= y \cdot (y \cdot 1) = y^2 \cdot 1 = 1 \cdot 1 = 1 \\ y \cdot 3 &= 3 \end{aligned}$$

So, here  $y$  would act as the permutation  $(1 \ 2)$ . If the other 2-cycles acted trivially, then this would be the only non-trivial permutation contradicting transitivity.

We conclude that at least two 2-cycles act non-trivially by permuting two elements. Since the 2-cycles generate  $S_3$ , we see that  $S_3$  **must** act by permutating  $U$ . Again, renaming elements if necessary, we may assume that the action of an element  $p \in S_3$  is

$$px = p(x)$$



**Note:** This discussion was necessary since we might as well have said that  $(1\ 2)$  maps  $1 \leftrightarrow 3$  and fixes 2, which would be very confusing, but logically possible.

Finally, consider the orbits. Orbits are calculated by taking an element and acting with all the elements of the group. We have

$$\begin{aligned} O_{(1,1)} &= \{g(1,1) \mid g \in S_3\} \\ &= \{(1,1), (2,2), (3,3)\} \\ O_{(1,2)} &= \{(1,2), (2,1), (1,3), (3,1), (2,3), (3,2)\} \end{aligned}$$

Thus we have a "diagonal" orbit and one other for the elements of the diagonal.

- [b] Suppose that  $S_3$  acts transitively on  $U$  but on  $V$  there are two orbits  $\{1\}$  and  $\{2, 3\}$ . As seen above, this means that  $S_3$  acts by permuting the elements of  $U$ , but in the case of  $V$  only one 2-cycle acts non-trivially e.g.  $(2\ 3)$ . Renaming elements if necessary, we have the orbits:

$$\begin{aligned} O_{(1,1)} &= \{(1,1), (2,1), (3,1)\} \\ O_{(2,1)} &= \{(2,3), (3,2), (1,2), (1,3), (2,2), (3,3)\} \end{aligned}$$

## Result

3 of 3

a) Since the actions are transitive on  $U, V$ , we arrive at two orbits.

b) Similar to a).

## 4. a

The stabilizer of a point  $p$  is the set of all symmetries  $s$  such that  $s(p) = p$ . In the top left image, we had translations and rotations by  $90^\circ$ . Obviously, translations don't fix any point. Rotations fix only the point about which they rotate. The rotations that preserve symmetry are those in the centers of the squares.

Put the origin in the center of one of the squares. Call the rotation by  $90^\circ$  around it  $\rho_p$ . The stabilizer of the origin is  $G_O = \{1, \rho_{pi/2}, \rho_{pi}, \rho_{3pi/2}\}$  and of some other point in the center of a square, it is

$$G_p = t_p G_O t_p^{-1} \cong C_4$$

$(t_p \rho_p t_p^{-1})$  is just a way to write a rotation around a point  $p$ . Thus the points with non-trivial stabilizers are the **centers** of the squares which are fixed by 4 rotations.

In the top right image we had rotational and bilateral symmetry. Rotations were again made around a point in the **center** of the squares in the pattern. If  $p$  is such a point, it will be preserved by rotations about it. Note that the reflections of the symmetry group don't pass through such points. Obviously, translations and glide symmetries don't preserve  $p$ . Therefore, as above

$$G_p \cong C_4$$

Consider the points which get preserved by horizontal reflection. These will be the **midpoints**  $m$  of the vertical edges of "vertical" rectangles. No other symmetry preserves this, so calling this reflection  $r$ , we have that

$$G_m = \{1, r\} \cong C_2$$

Similarly, the midpoints of horizontal edges of "horizontal" rectangles get preserved by reflection around a vertical axis through it.



In the lower left image, we also has rotational and bilateral symmetry. Similarly to the preceding, the points  $p$  in the **centers** of the squares, have stabilizer

$$G_p \cong C_4$$

since they are fixed only by the 4 possible rotations around the point (including the identity).

Lines of bilateral symmetry are given by horizontal and vertical lines through the center of the hexagrams. For the points  $p$  of the pattern, lying on these lines, we have that  $G_p \cong C_2$ .

The rest of the symmetries are translations or glide symmetries and thus don't have fixed points. Therefore, this checks all points.

## Step 4

4 of 5

In the lower right image, we had only translation and glide symmetry (by reflecting around a vertical line and then translating). None of these have any fixed points, and we conclude that all points have trivial stabilizers.

## Result

5 of 5

Top left - the centers of the squares have rotations as stabilizers, so  $G_p \cong C_4$ .

Top right - centers of squares have  $G_p \cong C_4$ , while midpoints of edges of rectangles are preserved by reflections, so  $G_m \cong C_2$ .

Lower left - the points in the centers of the squares have stabilizer  $G_p \cong C_4$ . For the points lying on the lines of horizontal/vertical reflection, only that reflection fixes them. Therefore  $G_p \cong C_2$  in that case.

Lower right - all symmetries are reflections and glide reflections. Therefore, no point is stabilized by a non-trivial element.

## 5. a

Let  $G$  be the group of symmetries of a cube, including the orientation-reversing symmetries. We can classify them as the orientation preserving:

- The identity, 1.
- Rotation by  $90^\circ$  about an axis going through the centers of opposing faces. Since we can rotate by  $\pm 90^\circ$  and there are three such axes, there are 6 such rotations.
- Rotation by  $180^\circ$  about by the above three axes - there are 3 such rotations.
- Rotation by  $180^\circ$  about an axis connecting the midpoints of two opposing edges. There are 6 such rotations.
- Rotation by an angle of  $120^\circ$  around a space diagonal. We can rotate by  $\pm 120^\circ$  around 4 such axes, for a total of 8 such rotations.

This gives all the 24 orientation-preserving symmetries. Consider a plane cutting the cube in half (any one will do). Reflecting the cube about that plane will also preserve the cube. Call that reflection  $r$  and let  $f$  be any other orientation-reversing symmetry. Then  $rf$  preserves orientation. Thus an arbitrary orientation-reversing symmetry  $f$ , can be written as

$$f = (fr)r$$

where  $fr$  preserves orientation. Thus, there are 24 more orientation-reversing symmetries, which are the results of composing the above symmetries by some reflection  $r$ .

## Result

2 of 2

We count the 24 orientation-preserving symmetries first. Composing any of these with a reflection will give 24 more orientation-reversing symmetries.

## 6. a

Let  $G$  be the group of symmetries of an equilateral triangular prism  $P$ , including the orientation-reversing symmetries. We count the possible elements of the group. We have

- the identity
- rotating by  $\pm \frac{2\pi}{3}$
- reflection about the plane intersecting the prism at middle height
- draw a vertical plane dividing a the prism in half. There are three possible ways to do this, and so 3 more reflections.
- in the plane in the middle of the prism, draw a line connecting opposing edges and midpoints of faces. Rotating by  $\pi$  around that axis is also a symmetry.
- the composition of reflection around the middle plane and rotation by  $\pm \frac{2\pi}{3}$ .

Consider now the symmetries fixing a given face of the prism  $P$ . Going by the list above

- the identity of course fixes the face
- rotating the prism doesn't fix any face
- reflection around the middle plane fixes any face
- reflection around the plane perpendicular to the face leaves the face unchanged
- rotation around lines in the middle fixes the plane if and only if the line is perpendicular to the face
- reflection composed with rotation doesn't preserve any face

Denoting the reflection around the middle plan by  $h$ , reflection around the vertical plane by  $v$  and rotation by  $\pi$  about the line in the center perpendicular to the face with  $\rho$ , we have the stabilizer group

$$G_F = \{1, h, v, \rho\}$$

Obviously, this group is not cyclic, and so is isomorphic to  $D_2$ .

### Result

3 of 3

There are 12 symmetries for a triangular prism. Only 4 of them fix a given face.

## 7. a

- [a] Let  $G = GL_n(V)$  operate on the set  $V = \mathbb{R}^n$  by left multiplication. The zero vector is obviously in its own orbit

$$G_0 = \{0\}$$

Now we consider some other orbit. Let  $x$  be a non-zero vector. Considering  $Ax$  for some matrices, it seems that we can get an arbitrary vector. Let  $y$  be another arbitrary non-zero vector. The question is can we always choose a matrix  $A$ , such that

$$Ax = y$$

for arbitrary non-zero  $x, y$ .

Since  $x, y$  are non-zero, they can be part of some basis. Let  $B_1 = \{x, \dots\}$  and  $B_2 = \{y, \dots\}$  two bases. Then the change of basis matrix  $P$  from  $B_1$  to  $B_2$  will be exactly such that  $Px = y$ . Change of basis matrices are always invertible, and therefore by choosing  $A = P$ , we have shown that  $Ax$  can be any vector. We conclude that all the non-zero vectors create the second orbit

$$G_x = \mathbb{R}^n \setminus \{0\}$$

- [b]) Let  $A = [a_{ij}]$  be an arbitrary invertible matrix. Then the condition

$$Ae_1 = e_1$$

written out leads to

$$Ae_1 = (a_{11}, a_{21}, \dots, a_{n1}) = (1, 0, \dots, 0) = e_1$$

and so it must be that  $a_{11} = 1$  and  $a_{k1} = 0$  for  $k > 1$ . This is also sufficient. Thus the stabilizer of  $e_1$  is the set of invertible matrices such that their first column contains a 1 as the first element, and the rest must be 0.

## Result

3 of 3

a) 0 is in its own orbit. Any non-zero vectors  $x, y$  are part of some basis. Choosing the appropriate change of basis matrix  $P$ , we have  $Px = y$ .  $P$  is always invertible, so the other orbit is  $\mathbb{R}^n \setminus \{0\}$ .

b) The first column of  $A \in G_{e_1}$  must be  $e_1$ .

## 8. a

Consider the set  $\mathbb{C}^{2 \times 2}$  of  $2 \times 2$  complex matrices. Let  $GL_2(\mathbb{C})$  act on it.

- [a]) Suppose the group acts by multiplication. It is easily seen that invertible matrices are in their own orbit, but it is not so clear about the non-invertible ones. Suppose  $X$  is a matrix and suppose  $A$  is an invertible matrix. Then if

$$AX = Y$$

the question is what we can say about  $Y$ ?

Notice that  $Yv = AXv = 0$  if and only if  $Xv = 0$ . Thus if two matrices  $X, Y$  are in the same orbit, we conclude that they must have the same kernels.

We want to prove the converse. Suppose that  $X, Y$  have the same kernels

$$\ker X = \ker Y$$

Let  $b$  be a basis for the kernel and let  $B \supseteq b$  be a basis for the whole space. Choosing the vectors in the appropriate order, we get a block matrix for  $X$  (or  $Y$ ) in the form

$$X = \begin{bmatrix} C_X & 0 \\ 0 & 0 \end{bmatrix}$$

(the right part of the matrix corresponding to the kernel). The non-zero columns must be linearly independent (else there would be another vector in the kernel of  $X$ ). We conclude that  $C_X$  is an invertible matrix. Similarly for  $Y$ , there is such a  $C_Y$ . If we let  $P$  be the change of basis matrix, then we have that

$$P \begin{bmatrix} C_Y C_X^{-1} & 0 \\ 0 & I \end{bmatrix} P^{-1} X = Y$$

and so  $X, Y$  are in the same orbit.

We conclude that the orbits are the sets of matrices with the same kernel.

- [b)] Suppose now that  $GL_2(\mathbb{C})$  acts with conjugation. By definition, this means that matrices are in the same orbit if and only if they are similar.

Over the complex numbers, every matrix is similar to some unique Jordan form matrix (up to order of Jordan blocks). These are  $2 \times 2$  matrices and we can just list their Jordan forms:

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix}$$

Any matrix is similar to one of these three types of matrices (of course, similar matrices must have the same eigenvalues). Thus there are infinitely many orbits of these types.

## Result

3 of 3

- a) We show that by choosing some good basis, there is an invertible matrix such that for  $X, Y$  with the same kernel, there is an invertible matrix such that  $AX = Y$ . On the other if  $AX = Y$ , then  $X, Y$  have same kernel. Thus the orbits are sets of matrices with the same kernel.
- b) The orbits are just similar matrices. We can compare to their Jordan normal forms. Thus matrices will be in the same orbit if and only if they have the same Jordan form.

## 9. a

- [a)] Let  $S$  be the set  $\mathbb{R}^{n \times m}$  of real  $m \times n$  matrices, and let  $G = GL_m(\mathbb{R}) \times GL_n(\mathbb{R})$ . The rule  $(P, Q) * A = PAQ^{-1}$  satisfies:

$$(I, I) * A = A$$

showing that the identity element of the group sends each element to itself. Next the associative law:

$$\begin{aligned} (P, Q) * ((R, S) * A) &= (P, Q) * (RAS^{-1}) \\ &= PRAS^{-1}Q^{-1} \\ &= (PR)A(QS)^{-1} \\ &= (PR, QS) * A \\ &= ((P, Q) \cdot (R, S)) * A \end{aligned}$$

We conclude that  $G$  operates on  $S$ .



- b) Let  $A$  be some matrix in  $S$  and consider its orbit by  $G$ . Recall the elementary row and column operations. These are invertible matrices. Recall that with the elementary operations, there was some composition of row operations  $P$  and column operations  $Q$  such that a matrix would reach the form

$$PAQ = I_r = \begin{bmatrix} 1 & & \\ & 1 & \\ & & \ddots \end{bmatrix}$$

where the matrix has  $r$  consecutive ones on the diagonal,  $r$  being equal to the rank of  $A$ . The rest of the elements are zeroes.

Thus all the elements of the same rank are contained in the same orbit.

Conversely, suppose two matrices  $A, B$  had the same rank. Then there are compositions of elementary matrices such that

$$PAQ = I_r = RBS$$

From which it follows that  $A = P^{-1}RBSQ^{-1} = (P^{-1}R, QS^{-1}) * B$ .

We conclude that the orbits are the sets of matrices with the same rank:

$$O_A = \{X \in S \mid \text{rank}(A) = \text{rank}(X)\}$$

The rank can have values  $0, 1, \dots, \min\{n, m\}$ . Thus there are  $\min\{n, m\} + 1$  orbits.

- [c] Assume that  $m \leq n$ . Consider the stabilizer of (the block matrix)  $[I \mid 0]$ . Notice that then

$$P[I \mid 0]Q^{-1} = [P \mid 0]Q^{-1}$$

Write  $Q^{-1}$  as a block matrix,

$$Q^{-1} = \begin{bmatrix} Q_m & Q'_m \\ X & Y \end{bmatrix}$$

where  $Q_m$  is of size  $m \times m$  and the other matrices are of size appropriate to that. Now multiply the blocks:

$$[P \mid 0] \begin{bmatrix} Q_m & Q'_m \\ X & Y \end{bmatrix} = [PQ_m \mid PQ'_m]$$

So if  $(P, Q)$  is in the stabilizer, this must be  $[I \mid 0]$ . So we must have that  $Q_m = P^{-1}$  and  $PQ'_m = 0$ . Since  $P$  is invertible this means that  $Q'_m = 0$ .

We conclude that the stabilizer of  $[I \mid 0]$  is made of  $(P, Q)$  where  $P$  is an arbitrary invertible matrix and  $Q^{-1}$  is of the form

$$\begin{bmatrix} P^{-1} & 0 \\ X & Y \end{bmatrix}$$

where  $X, Y$  are some matrices ( $Y$  must be invertible).

## Result

4 of 4

- Check the that the identity element preserves elements and the associative law.
- Elementary operations can transform any matrix of the same rank one into another. Thus the orbits are given by matrices of the same rank.
- Note that  $P[I \mid 0] = [P \mid 0]$ . Then multiplying by  $Q^{-1}$ , we see that  $Q^{-1}$  must have in the top left corner  $P^{-1}$  and 0 in the top right corner.

- [a)] Consider the matrix

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

and suppose that  $GL_2(\mathbb{R})$  acts on  $2 \times 2$  matrices. We find the orbit and stabilizer of this matrix.

Consider first its orbit. This is a diagonal matrix. Let  $A$  be any matrix with eigenvalues 1, 2. Since these are real eigenvalues, we know that there is a real matrix  $P$  such that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Also, we can reverse the conjugation to recover  $A$ . We conclude that the orbit of the matrix are the real matrices with eigenvalues 1, 2.

Now, suppose  $P$  is in the stabilizer of the matrix above. Then we want to have  $P^{-1}XP = X$  or equivalently  $XP = PX$ , so  $P$  must commute.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\begin{bmatrix} a & 2b \\ c & 2d \end{bmatrix} = \begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix}$$

We see that we must have  $b = c = 0$ , but  $a, d$  can be arbitrary. We conclude that the stabilizer subgroup is the group of diagonal matrices.

- [b)] The orbit of the matrix is equal to all the possible values of  $AX$  where  $A$  is a matrix in  $G = GL_2(\mathbb{F}_5)$ . There is a bijection between the orbit  $O_A$  and the quotient  $G/G_A$ . Just as above, the stabilizer is given by diagonal matrices (invertible) matrices. There  $4 \times 4 = 16$  of those.

The order of the group is given by  $|G| = (5^2 - 1)(5^2 - 5) = 480$  (this can be counted by considering the fact the columns must be linearly independent). Thus

$$|O_A| = |G/G_A| = |G|/|G_A| = 480/16 = 30$$

## Result

3 of 3

a) The orbit is given by matrices with eigenvalues 1, 2. The stabilizer is given by diagonal matrices.

b) Use the orbit stabilizer theorem. Just as in a), the stabilizer is the diagonal matrices - there are  $4 \times 4 = 16$  of those. Thus the orbit has order  $480/16 = 30$ .

11. a



We first discuss some basics about permutations.

If  $g$  is a permutation, then conjugation acts upon a cycle as

$$g(a_1 a_2 \dots a_k)g^{-1} = (g(a_1) g(a_2) \dots g(a_k))$$

Proof: We apply both sides to some element of  $S_n$ . If  $a$  is not among  $g(a_1), \dots, g(a_k)$  both the left and right side simply fix it. And if  $a$  is among them, it is to check that both sides permute it by one index up to  $g(a_{i+1})$ .

The conclusion is that

**conjugates of  $k$ -cycles are again  $k$ -cycles**

. Writing a permutation as a product of distinct permutations, we find that

**conjugate permutations have the same cycle structure.**

Conversely, if two permutations have the same cycle structure, it is not hard to choose a  $g$  as above such that they are conjugate.

Consider the orbits under conjugation. These are

- the identity
- six 2-cycles  $(a b)$
- eight 3-cycles  $(a b c)$
- six 4-cycles  $(a b c d)$
- three  $2 \times 2$ -cycles  $(a b)(c d)$

Now, if a  $H$  is a normal subgroup and  $h \in H$ , then also  $ghg^{-1}$  for **any** other elements of  $h$ . Thus if  $h$  is in  $H$ , so is also the orbit  $O_h$  (of conjugates of  $h$ ). We see that the orbits (also called *conjugacy classes* have orders

$$1, 6, 8, 6, 3$$

Since  $H$  must contain the identity and be of order 12, the only possibility is that it contains the orbit of 3-cycles and  $2 \times 2$  cycles. But then  $H$  is the subgroup of even permutations, that is,  $A_4$ .

## Result

3 of 3

Note that conjugate permutations have the same cycle structure. Then acting with conjugation on  $S_4$ , we have 5 orbits (i.e. conjugacy classes) of sizes 1, 6, 8, 6, 3.

If  $H$  is a normal subgroup, then if  $h \in H$  so also  $ghg^{-1} \in H$ . Thus a normal subgroup cannot contain just parts of an orbit.  $H$  must contain the identity and this leaves us with 11 more elements. The only possibility is that  $H = A_4$ .

## Section 8

1. a

Consider the rule

$$P * A = PAP^t$$

where  $P \in GL_n$  and  $A$  is an  $n \times n$  matrix. We prove that this is an operation on the set of all matrices. First, for  $P = I$ , we have

$$I * A = IAI^t = A$$

and for the associative law:

$$\begin{aligned} (PQ) * A &= PQ A (PQ)^t \\ &= PQ A Q^t P^t \\ &= P(Q * A)P^t \\ &= P * (Q * A) \end{aligned}$$

We conclude that the rule indeed defines an operation of  $GL_n$  on the set of all matrices.

## Result

2 of 2

Check the two basic requirements for a group operation.

### 2. a

Let  $G$  operate on  $G/H$ . We find the stabilizer of the coset  $[aH]$ . These are the elements of  $G$  such that

$$g[aH] = [gaH] = [aH]$$

Thus  $aH$  and  $gaH$  must be the same cosets. So

$$\begin{aligned} gaH &= aH & / a^{-1}. \\ a^{-1}gaH &= H \end{aligned}$$

This is possible if and only if  $a^{-1}ga \in H$  or equivalently  $g \in aHa^{-1}$ . Thus the stabilizer subgroup of  $[aH]$  is

$$G_{[aH]} = aHa^{-1}$$

## Result

2 of 2

Note that the stabilizer is the set of  $g \in G$  such that  $gaH = aH$ . This is equivalent to  $a^{-1}gaH = H$  showing that  $aHa^{-1}$  is the stabilizer.

### 3. a

Consider the set of left cosets of the stabilizer subgroup of a vertex. A vertex  $v$  is stabilized only by the reflection  $y$  through that vertex. So the stabilizer subgroup of a vertex  $v$  is

$$H = \{1, y\}$$

Let  $x$  be rotation by  $90^\circ$ . The cosets are

$$\{1, y\}, \{x, xy\}, \{x^2, x^2y\}, \{x^3, x^3y\}$$

Obviously,  $D_4$  acts transitively on the set of vertices  $S$  (i.e. any vertex can be derived from another by some operation of  $D_4$ ). Therefore the orbit of any vertex is the whole set  $S = \{v, v_1, v_2, v_3\}$ , where the vertices are in counter-clockwise order  $v, v_1, v_2, v_3$ . The bijection between  $G/H$  and  $O_v$  is explicitly given by

$$\begin{aligned}\{1, x\} &\leftrightarrow v \\ \{x, xy\} &\leftrightarrow xv = v_1 \\ \{x^2, x^2y\} &\leftrightarrow x^2v = v_2 \\ \{x^3, x^3y\} &\leftrightarrow x^3v = v_3\end{aligned}$$

## Result

2 of 2

A vertex  $v$  is stabilized by a reflection about a line through it. Thus there are 4 cosets of the stabilizer group and the orbit is of size 4. The bijection is found by inspection.

## 4. a

Let  $H$  be the stabilizer of the index 1 for the operation of the symmetric group  $G = S_n$  on the set of indices  $\{1, 2, \dots, n\}$ .

$H$  is the set of all permutations fixing 1 - since the rest of the  $n - 1$  indices can be permuted arbitrarily, we conclude that  $H$  is isomorphic to  $S_{n-1}$  and is of size  $(n - 1)!$ . The left cosets of  $H$  are

$$\begin{aligned}H &= \{\phi \in S_n \mid \phi(1) = 1\} \\ (1\ 2)H &= \{\phi \in S_n \mid \phi(1) = 2\} \\ (1\ 3)H &= \{\phi \in S_n \mid \phi(1) = 3\} \\ &\dots \\ (1\ n)H &= \{\phi \in S_n \mid \phi(1) = n\}\end{aligned}$$

This corresponds to the elements of the orbit of 1 - which is simply the whole set  $\{1, 2, \dots, n\}$ . Thus we have the bijection

$$\begin{aligned}H &\leftrightarrow 1 \\ (1\ 2)H &\leftrightarrow (1\ 2)1 = 2 \\ (1\ 3)H &\leftrightarrow (1\ 3)1 = 3 \\ &\dots \\ (1\ n)H &\leftrightarrow (1\ n)1 = n\end{aligned}$$

## Result

2 of 2

Note that the stabilizer is equal to all the permutations fixing 1. There is only one orbit of 1 (the whole set of indices). There is a bijection between the  $n$  cosets and  $n$  indices, in the form of  $(1\ a)H \leftrightarrow a$ .

## Section 9

### 1. a

We use the counting formula to determine the orders of the groups of rotational symmetries of a cube and of a tetrahedron.

Take the cube first. Every face is symmetrical being rotated by  $90^\circ$  and multiples of that. This means that there are 4 rotations in the stabilizer of a face. A cube has 6 faces, and so the rotational symmetry of a cube has

$$4 \cdot 6 = 24$$

elements.

### Step 2

2 of 3

Take the tetrahedron now. It has 4 triangles as faces. There are 3 rotations fixing each face. We conclude that there

$$3 \cdot 4$$

rotational symmetries of the tetrahedron.

### Result

3 of 3

Multiply the number of faces and the number of rotational symmetries. A cube has squares as faces, and so has 4 rotations fixing it, while a tetrahedron's face has 3.

This gives  $4 \cdot 6 = 24$  rotations for the square and  $3 \cdot 4$  for the tetrahedron.

### 2. a

Let  $G$  be the group of rotational symmetries of a cube, let  $G_v, G_e, G_f$  be the stabilizers of a vertex  $v$ , an edge  $e$ , and a face  $f$  of the cube, and let  $V, E, F$  be the sets of vertices, edges, and faces, respectively.

We first describe the subgroups mentioned above.

- The stabilizer of  $v$  must contain only rotations which fix a vertex. The only lines of rotation through vertices are the space diagonals. There is only one diagonal passing through a vertex, thus  $|G_v| = 3$  and contains the identity and two possible rotations by  $\frac{2\pi}{3}$  around a space diagonal.
- The stabilizer of  $e$  must contain only rotations which fix an edge. This can only happen for the rotation about the line connecting opposite edges at their midpoints. Thus  $|G_e| = 2$  and contains the identity and rotation around the line connecting opposite midpoints of edges.
- The stabilizer of  $f$  must contain only rotations which fix a face. This happens for a rotation about an axis that passes through the center of that face. These are rotations by  $\frac{2k\pi}{4}$ , and so  $|G_f| = 4$ .

Consider now the subgroup  $G_v$  of rotations around a space diagonal. This fixes the vertex on end of the diagonal, and vertex on the other. The vertices adjacent to one of these get rotated one into another. Thus the decomposition of the sets of vertices by  $G_v$  is

$$|V| = 1 + 3 + 3 + 1$$

Consider now the edges. The orbits must be of size dividing  $|G_v| = 3$ . Since all of the edges move when rotated around space diagonal, we conclude that they are all in orbits of size 3. Thus

$$|E| = 3 + 3 + 3 + 3$$

Faces adjacent to the vertices are rotated around the space diagonal one into another. Thus

$$|F| = 3 + 3$$

Consider the subgroup  $G_e$  which has the rotation by  $180^\circ$  about the axis connecting opposite midpoints. The rotation is of order 2, so if it doesn't fix an element, the element has an orbit of size 2. Obviously, this rotation doesn't fix any vertex, thus

$$|V| = 2 + 2 + 2 + 2$$

Consider now the edges. The only edges fixed are the ones which the axes intersect. All the others are moved, so

$$|E| = 1 + 1 + 2 + 2 + 2 + 2$$

No face is fixed, thus they are all in orbits of size 2, so

$$|F| = 2 + 2 + 2$$

Consider the subgroup  $G_f$  which has the rotations by  $90^\circ$  about the axis connecting opposite faces. The group is of order 4, so if it doesn't fix an element, the element has an orbit of size 2 or 4. Vertices of the opposing faces are rotated one into another so

$$|V| = 4 + 4$$

Consider now the edges. The edges also rotate cyclically by  $90^\circ$  one into another. Thus they have orbits of size 4 and we get the decomposition:

$$|E| = 4 + 4 + 4$$

The faces through which the axis passes stay fixed, whereas the other faces rotate cyclically one into another, thus they are in an orbit of size 4. The decomposition is:

$$|F| = 1 + 1 + 4$$

## Result

5 of 5

$G_v$  is the group of rotations by  $2\pi/3$  about the space diagonal.  $G_e$  is the group of rotating by  $\pi$  around the axis connecting opposing midpoint.  $G_f$  rotates the cube by  $90^\circ$ . Since the order of an orbit must divide the order of the group, it is not hard to establish the orbits.

3. a



Let  $G$  be the group of symmetries of the dodecahedron. Suppose it operates on the faces. Let  $f$  be some face. If  $O_f$  is the orbit of  $f$  and  $G_f$  its stabilizer, we have

$$|G| = |O_f| |G_f|$$

There is only one orbit of the faces (since each one can be rotated in another), thus  $|O_f| = 12$  since there are 12 faces. We only need calculate the stabilizer. The dodecahedron has pentagonal faces, so they have rotational symmetry, rotating by  $\frac{2\pi}{5}k$ .

One can also connect a line between vertices and opposing lines in the midpoint. Reflecting the pentagon about that line is a symmetry of the pentagon. Similarly, drawing a plane through that line which halves the pentagon is a symmetry of the pentagon.

There can be no translational symmetries. Thus, there are 10 symmetries which fix the pentagonal faces. By the above formula, we have

$$|G| = 12 \cdot 10 = 120$$

## Result

2 of 2

Every face is preserved by 10 symmetries - 5 rotations (including the identity) and 5 reflections. Since there are 12 faces, we have that the group of symmetries is  $12 \cdot 10 = 120$ .

## 4. a

Let  $T'$  be the group of all symmetries of a regular tetrahedron, including orientation-reversing symmetries.

The tetrahedron has 4 vertices.  $T'$  operates on them by permuting them. Therefore, there must be a homomorphism

$$T' \rightarrow S_4$$

mapping the elements of  $T'$  to corresponding permutations. Let us calculate first the order of  $T'$  which should tell us something.

Every face has triangular symmetries, the identity, two rotations by  $\pi/3$  and three reflections. Thus the stabilizer of any face has order 6. Since there are 4 faces, we have that the group of symmetries is of size

$$6 \cdot 4 = 24$$

Now, note that every symmetry of the tetrahedron permutes the vertices differently (e.g. the eight rotation by  $120^\circ$  around the vertices do so and the three rotations by  $180^\circ$  around opposing edges do so - since these do, so do reflections). We conclude that the homomorphism between  $T'$  and  $S_4$  is injective and since both groups have 24 elements, we conclude that  $T' \cong S_4$ .

## Result

2 of 2

By the counting formula, we find that  $|T'| = 4 \cdot 6 = 24$ . Considering the symmetries, we see that all permute the vertices differently. Therefore,  $T'$  corresponds to the symmetric group on 4 elements.

## 5. a



Symmetry is defined as the mode by which the shape becomes exactly like the one when it is moved in some way.

[Comment](#)

## Step 2 of 4 ^

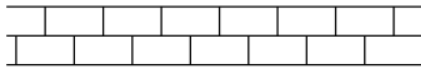
Let  $J$  be a section of an  $I$ -beam, which one can think of as the product set of  $I$  and the unit interval.

To identify: Its group of symmetries, orientation-reversing symmetries induced

The identity operation is denoted as the letter  $I$

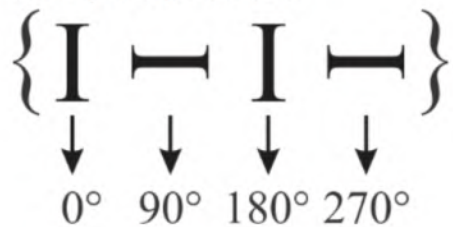
Since, the letter  $I$  looks like the number 1 and if the number 1 is multiplied by any number then the solution remains unchanged.

Sometimes, the symmetry as a combination of translational and mirror symmetry as is of the form of a brick as shown below;



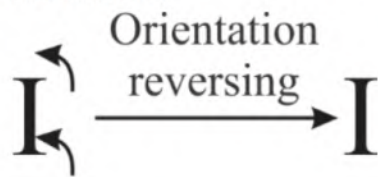
Now, if the letter  $I$  is multiplied with the unit interval then the product will be the letter  $I$  itself.

Thus, the group of symmetries will be;



Further the orientation-reversing symmetry is defined as the reflectional symmetry and the rotational symmetry that shares the same point of symmetry.

Now, the orientation reversing symmetry of the letter  $I$  means that changing the face of the letter that is shown below;

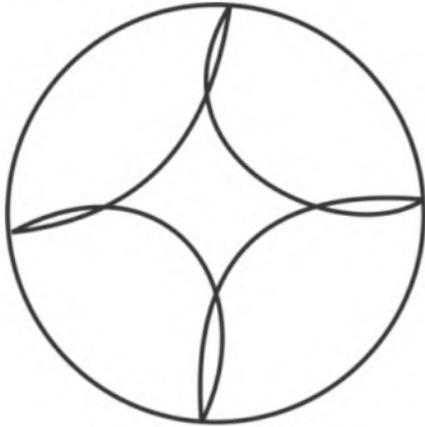


6. a

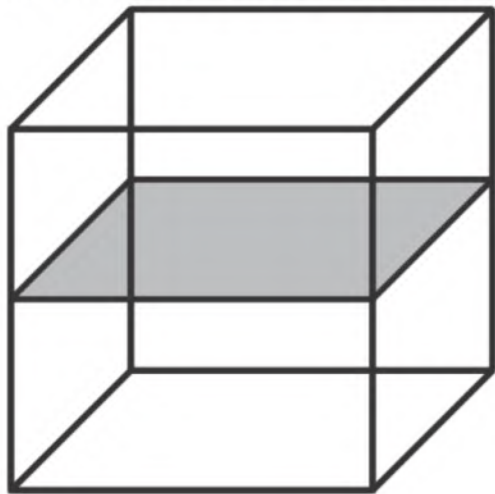
Symmetry defines that one shape becomes exactly like another when it is cut or moved along any line

First note that the baseball is the same as a cube with an edge path representing the seam as given below;

Consider the baseball as given below:



And, its corresponding cube with an side path representing the given seam is given below:



Now, consider the square  $T$  in the middle of the cube considering with each of its vertices on the path representing the given seam.

The vertices represent the four non-changing points of the seam of the given round shaped or the spherical shaped baseball.

Let the symmetry be denoted by  $T$

So, all the symmetries should be preserved by  $T$

Any one side of the cube with having two edges of the seam should be sent to a seam of side which is only having two edges.

That is all the symmetry should maintain the required vertical fibers of the cube this means that the top of the cube can only be sent to the top or bottom, so that the middle part will be fixed.

This tells about the homomorphism of the symmetry group of the baseball with the symmetry group of square, which is the Dihedral group  $D_4$

And, since the above proved homomorphism is one to one all the symmetries of the baseball will be obtained by the action on the symmetries, that is  $g$ .

Hence, the required result is  $D_4$

## Section 10

### 1. a

Consider the set  $S$  of subsets of order of  $D_3$ . Let  $x, y$  be the generators of  $D_3$ . The set  $S$  has order

$$\binom{|D_3|}{3} = \binom{6}{3} = 20$$

To calculate the orbits, choose any set and multiply with all possible elements of  $D_3$ . Let  $O$  be an orbit and  $H$  be the stabilizer of a subset  $S$ . Recall that

$$|D_6| = 6 = |O||H|$$

Thus to find the size of the orbits, it suffices to find the orders of their stabilizers. Consider the subset

$$\{1, x, x^2\}$$

It is immediately seen that it is fixed by  $1, x, x^2$  and not fixed by multiplication with the other elements. Thus the stabilizer is of order 3, and so

$$|O_{\{1, x, x^2\}}| = 6/3 = 2$$

Consider now some other subset of the form

$$\{1, x, a\}$$

where  $a = y, yx, yx^2$ . We find the stabilizer. Suppose that  $b$  stabilizes this set i.e.

$$b \cdot \{1, x, a\} = \{b, bx, ba\}$$

These two sets should be equal. The right hand side set must have equal members to the left. If  $b = 1$ , that is fine. Now, if  $b = x$ , then the right side contains  $x^2$  and the left doesn't. Finally, if  $b = a$ , then the right side contains  $ax$  and the left doesn't.

We conclude that the only element stabilizing such a subset is 1. Therefore the orbit has size

$$|O_{\{1, x, a\}}| = 6/1 = 6$$

Finally, note that the orbits of  $\{1, x, a\}$  for distinct  $a$  are disjoint. Indeed, if one was to have

$$b \cdot \{1, x, a\} = \{1, x, a'\}$$

Then  $b$  would have to be either  $1, x, a'$  but none of that works. We conclude that the orbits of  $\{1, x, a\}$  where  $a = y, yx, yx^2$  are disjoint.

This gives us one orbit of size 2 and three of size 6. Since the set in total has  $20 = 2 + 6 + 6 + 6$  elements, there are no other orbits.

### Result

4 of 4

There are  $\binom{6}{3} = 20$  subsets of order 3 of  $D_3$ . Letting  $x, y$  be the generators of  $D_3$ , we consider the orbits of  $\{1, x, a\}$  where  $a = x^2, y, yx, yx^2$ . For  $x^2$  the orbit has order 2, whereas for the rest it has order 6. The orbits of these elements are distinct, so this makes all of them.

### 2. a

Let  $S$  be a finite set on which a group  $G$  operates transitively, and let  $U$  be a subset of  $S$ . Let  $g \in G$  be arbitrary and let  $A_x$  be all the subsets  $gU$  which contain the element  $x$ , so

$$A_x = \{gU \mid g \in G, x \in gU\}$$

Now, since  $G$  acts transitively, for any other element  $y$ , there exists an element  $z$  such that

$$z \cdot x = y$$

We define  $f : A_x \rightarrow A_y$  by  $f(gU) = zgU$ . Since  $x \in gU$ , then  $y = zx \in zgU$ . That this is a bijection is obvious since it is an invertible function:

$$f^{-1}(gU) = z^{-1}gU$$

Therefore there is a bijection between the sets containing any two elements. We conclude that all elements appear in evenly many sets.

## Result

2 of 2

Since  $G$  operates transitively, there is a  $z$  such that  $z \cdot x = y$  for any  $x, y$ . Then if  $gU$  is some subset of  $S$  containing  $x$ , then  $zgU$  is a set containing  $y$ . Since multiplication by  $z$  is a bijection, the number of such sets is the same.

## 3. a

Consider the operation of left multiplication by  $G$  on the set of its subsets. Let  $U$  be a subset such that the sets  $gU$  partition  $G$ . Let  $H$  be the unique subset in this orbit that contains 1. We show that  $H$  is equal to its stabilizer subgroup  $G_H$ .

Suppose  $g \in G_H$ . This means that

$$gH = H$$

since  $H$  contains 1, the left hand side contains  $g$ , showing that  $g \in H$ .

Now, assume  $g \in H$ . Then

$$gH$$

contains  $g$  again since  $1 \in H$ . But the sets  $gU$  partition  $G$ , meaning they are either equal or disjoint. Thus, if  $gH$  has an element in common with  $H$ , it must be  $H$ . We conclude that

$$gH = H$$

and so  $g \in G_H$ .

We see that  $H$  is equal to its stabilizer subgroup  $G_H$  and so is a subgroup.

## Result

2 of 2

Show that  $H = G_H$ , where  $G_H$  is the stabilizer subgroup of  $H$  under left multiplication. The equality of sets is checked by proving two inclusions.

# Section 11

## 1. a

Let  $G$  be a group operating on a set  $S = \{a, b, c, d\}$ . There is a bijection between operations of  $G$  on  $S$  and homomorphisms  $G \rightarrow \text{Perm}(S)$ . Let  $G = S_3$ , and suppose  $S$  is a four element set. Then the possible ways in which  $S_3$  operates correspond to homomorphisms

$$\varphi : S_3 \rightarrow S_4$$

The image of  $S_3$  must be isomorphic to some subgroup of  $S_4$ . We have the possibilities

- $\varphi(S_3) \cong S_3$  - this means that  $\varphi(S_3)$  must be a subgroup of order 6 in  $S_4$ . Being of order 6, it must contain an element of order 2 and 3. There are two possibilities:
- $\varphi(S_3)$  contains a 2-cycle and a 3-cycle

$$(a\ b), (a\ c\ d) \in \varphi(S_3)$$

(note that since there are 4 elements, they must have one element in common). Now, suppose  $b \neq c, d$ . But then

$$(a\ b)(a\ c\ d) = (a\ c\ d\ b)$$

would be of order 4 in a group of 6 elements. We conclude that  $b$  must be one of  $c, d$  i.e.  $\varphi(S_3)$  permutes three elements.

- $\varphi(S_3)$  contains a  $2 \times 2$ -cycle and a 3-cycle

$$(a\ b)(c\ d), (a\ b\ c)$$

Then

$$(a\ b\ c)(a\ b)(c\ d) = (a\ c\ d)$$

which is impossible since then  $\varphi(S_3)$  would contain two distinct 3-cycles which generate a group larger than size 6.

We see that the only possibility is that  $S_3$  permutes three elements.

- $\varphi(S_3) \cong C_3$  - but since  $G/\ker \varphi \cong \varphi(G)$  by the first isomorphism theorem, this would mean that  $\varphi$  has kernel of size 2. The kernel is always a normal subgroup, but  $S_3$  does not have normal subgroups of order 2. We conclude that this case is impossible.
- $\varphi(S_3) \cong C_2$  - this is possible since  $\ker \varphi$  of order 3 is a normal subgroup. All subgroups of order 2 are isomorphic, so  $\varphi(S_3)$  contains the identity and an arbitrary element of order 2. There are two possibilities:

$$(a\ b) \quad \text{or} \quad (a\ b)(c\ d)$$

Thus either  $S_3$  permutes two elements, or permutes two pairs of elements.

- Finally,  $\varphi(S_3) = \{1\}$  and then  $S_3$  acts trivially on  $S$ .

## Result

3 of 2

Recall that the possible group operations of  $G = S_3$  on  $S$  correspond to the possible homomorphisms  $\varphi$  of  $G$  into  $\text{Perm}(S)$ .  $\varphi(G)$  can be of size 1, 2, 6 and one analyzes each case separately.

## 2. a



Let  $G = T$  be the tetrahedral group.  $T$  consists of 12 rotations, eight which rotate the tetrahedron by  $120^\circ$  around the vertices, and three which rotate the tetrahedron by  $180^\circ$  around the axis connecting midpoints of opposing edges.

Let  $S$  be a set with two elements and suppose  $T$  operates on  $S$ . The possible operations correspond to homomorphisms

$$\varphi : T \rightarrow S_2$$

There are two possibilities

- Suppose that  $\varphi(T) \cong S_2$ . This would mean that  $|\ker \varphi| = 12/2 = 6$  (by the first isomorphism theorem). This would mean that  $T$  has  $H = \ker \varphi$  as a subgroup of size 6. We show that this is impossible.

Suppose to the contrary that there exists an  $H$  such that  $|H| = 6$ . Note that  $H$  cannot contain only rotations by  $180^\circ$  since there are only 3 of those. Also it cannot contain only rotations by  $120^\circ$ , since those can be composed with one another to get the rest. We conclude that  $H$  must contain a  $180^\circ$  and a  $120^\circ$  rotation. But the composition of such two will be a  $180^\circ$  around another vertex. Thus the two elements generate all of  $T$ . Thus necessarily  $|H| = 12$ , a contradiction.

- Suppose  $\varphi(T) \cong \{1\}$ . This means that the action corresponds to the trivial action.

We see that  $T$  can only act in a trivial way on a set of two elements.

## Result

2 of 2

The possible operations of  $T$  on a set of two elements correspond to the possible homomorphisms of  $T$  into  $S_2$ . We show that only the trivial homomorphism is possible. Thus  $T$  operates trivially.

## 3. a

Let  $S$  be a set on which a group  $G$  operates, and let  $H$  be the subset of elements  $h$  such that  $hs = s$  for all  $s$  in  $S$ . Thus  $H = G_S$ , the stabilizer of the whole set.

Let  $g \in G$  and  $h \in H$  be arbitrary. This means that  $hs = s$  and

$$(g^{-1}hg)s = g^{-1}(h(gs))$$

Now,  $gs$  is an arbitrary element of  $S$ , and so  $h(gs) = gs$ . Thus

$$\begin{aligned} (g^{-1}hg)s &= g^{-1}(h(gs)) \\ &= g^{-1}(gs) \\ &= s \end{aligned}$$

This shows that  $g^{-1}hg \in H$ . We conclude that  $H$  is a normal subgroup of  $G$ .

## Result

2 of 2

If  $hs = s$  for all  $s$ , then  $(g^{-1}hg)s = g^{-1}(h(gs)) = (g^{-1}g)s = s$ . This shows that  $H$  is normal.

## 4. a



Let  $G$  be the dihedral group of  $D_4$  of symmetries of a square. Consider the action of  $G$  on the vertices. The action is faithful if there is no element which acts as the identity (besides the identity itself).

Certainly, (non-trivial) rotations fix no vertex. Also no reflection fixes all vertices. Therefore the action is faithful.

Consider now the action on the two diagonals. They are not fixed by rotations by  $90^\circ$  but **are** fixed by rotations by  $180^\circ$ . (Also reflections around the diagonals themselves fix them.) We conclude that the group doesn't act faithfully.

## Result

2 of 2

It is easy to see that rotations permute some vertices, and reflections do just as well. Therefore, no element of  $D_4$  acts as the identity on the vertices, and so this action is faithful.

On the other hand, diagonals are fixed by rotation by  $180^\circ$ , and therefore  $D_4$  doesn't act faithfully.

## 5. a

Suppose a group  $G$  operates faithfully on a set  $S$  of five elements, and there are two orbits, one of order 3 and one of order 2. Recall that every group operation corresponds to a homomorphism  $\varphi(g) = m_g$ , where  $m_g$  is left multiplication:

$$\varphi : G \rightarrow \text{Perm}(S)$$

Now, since there are 2 orbits, the elements in the orbits can permute only one into another. Thus the even stronger statement holds that

$$\varphi : G \rightarrow S_3 \times S_2$$

Since  $G$  operates faithfully, we must have  $\ker \varphi = \{1\}$ . Thus

$$G \cong G / \ker \varphi = \varphi(G)$$

and so  $G$  must be isomorphic to a subgroup of  $S_3 \times S_2$ .

Recall the fact that the size of an orbit must divide the order of the group. Since we have orbits of size 2 and 3, we conclude that

$$6 \mid |G|$$

Since  $G$  is a subgroup of  $S_3 \times S_2$  of order 12, we conclude that  $G$  could either be of size 6 and then it has to be isomorphic to  $C_6$  or  $S_3$ , or if  $G$  is of order 12, then  $G = S_3 \times S_2$ . We show that all cases are possible:

- It is possible that  $G = S_3$ . Let  $G$  naturally permute the 3 elements in the orbit of size 3. For the other two elements, we can have all elements act trivially except some 2-cycle.

By its very construction, this group operation satisfies the requirements.

- It is possible that  $G = C_6$ . Suppose  $z$  is its generator. Let  $G$  act on the orbit of size 3 by permuting the elements cyclically. Let  $G$  act non-trivially on the orbit of size 2. (This is possible since  $C_6$  is abelian, so there is always a homomorphism to subgroups of any size dividing  $|G|$ ). Explicitly, this can be stated by

$$zx = \begin{cases} z^2x & x \in O_3 \\ z^3x & x \in O_2 \end{cases}$$

- Finally, if  $G = S_3 \times S_2$ , then let this group act on  $S$  by

$$(\sigma, \tau) \cdot x = \begin{cases} \sigma x & x \in O_3 \\ \tau x & x \in O_2 \end{cases}$$

where  $O_3, O_2$  are the orbits of size 3, 2.

## Result

3 of 3

There is always a homomorphism from  $G$  to  $\text{Perm}(S)$ , the group of permutations of the set  $S$ . Here we can say even more that it must map to  $S_3 \times S_2$ . Having orbits 2, 3, it must be that  $6 \mid |G|$ . There are three possible subgroups of  $S_3 \times S_2$  satisfying this. Are all possible.

## 6. a

Let  $F = \mathbb{F}_3$ . Consider the vector space  $F^2$ . The one-dimensional subspaces are

$$\begin{aligned} V_1 &= \{(0, 0), (1, 1), (2, 2)\} \\ V_2 &= \{(0, 0), (1, 2), (2, 1)\} \\ V_3 &= \{(0, 0), (0, 1), (0, 2)\} \\ V_4 &= \{(0, 0), (1, 0), (2, 0)\} \end{aligned}$$

The group  $GL_2(F)$  acts on the set of these four vector spaces by left multiplication. Therefore, there is a corresponding homomorphism  $\varphi(A) = m_A$ , since  $m_A$  can be considered a permutation of the four vector spaces.

$$\varphi : GL_2(F) \rightarrow S_4$$

Consider the kernel of this map. These are the invertible matrices  $A$ , such that

$$\varphi(A) = m_A = 1$$

This happens if and only if  $A$  doesn't permute any of the spaces above. This would mean that

$$\begin{aligned} A(1, 0)^t &= (\pm 1, 0)^t \\ A(0, 1)^t &= (0, \pm 1)^t \end{aligned}$$

(with  $-1 = 2$  in  $\mathbb{F}_3$ ). Writing out  $A$  with some elements, we have that  $A$  is in the kernel if

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} \pm 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix} \end{aligned}$$

This leads to  $b = c = 0$  and  $a, d$  being  $\pm 1$ . Now, if  $a \neq d$ , then one checks that  $A(1, 1)^t \neq (\pm 1, \pm 1)$  as it should if  $A$  was in the kernel.

We conclude that  $a = d = \pm 1$ , and so the kernel consists of the matrices  $\pm I$ .

Since the kernel has order 2, we have that

$$|\varphi(GL_2(F))| = |GL_2(F)/\ker \varphi| = 48/2 = 24$$

(The order of  $GL_2(F)$  is calculated by counting how many linearly independent columns there can be - choose the first column in 8 ways - for each of those there are six columns independent from it. Thus there are  $8 \times 6 = 48$  possible choices.)

We see that the image has the same size as  $S_4$  and so

$$\varphi(GL_2(F)) \cong S_4$$

## Result

$I_4$  of  $I_4$

The kernel of  $\varphi$  are those matrices that fix all the spaces. For  $A$  to be in the kernel, we find that  $A$  must be  $\pm I$ . The group  $GL_2(F)$  has size  $8 \cdot 6 = 48$ . Therefore the image of  $\varphi(GL_2(F))$  has size  $48/2 = 24$  and so must be isomorphic to  $S_4$ .

7. a

- [a]) Assume that  $D_4$  acts faithfully on a set of  $n$  elements. As noted above,  $|D_4| = 8$  must divide the order of  $S_n$ . This can happen first for  $n = 4$ . And indeed, taking e.g. the vertices of a square,  $|D_4|$  acts faithfully on them.

Thus  $D_4$  can act faithfully on a set of order minimally 4.

## Step 2

2 of 4

- [b]) Assume that  $D_6$  acts faithfully on a set of  $n$  elements. As noted above,  $|D_6| = 12$  must divide the order of  $S_n$ . This can happen first for  $n = 4$ . But  $D_6$  has an element of order 6 which is certainly impossible for  $S_4$ . This is also impossible for  $S_5$ .

For  $S_6$  it is possible and in fact occurs. One simply considers the action of  $D_6$  on a hexagon. It permutes the vertices and does so faithfully.

Thus  $D_6$  can act faithfully on a set of order minimally 6.

- [c]) Let  $H$  be the quaternion subgroup and let it act faithfully on a set of  $n$  elements. This means that  $|H|$  must divide  $|S_n|$ . This can happen first for  $n = 4$ . We will show that it happens first for  $n = 8$ . Recall that the quaternion group has distinct elements  $i, j, k$  such that

$$\begin{aligned} i^2 &= j^2 = -1 \\ ij &= -ji = j^{-1}i \end{aligned}$$

and so  $iji^{-1} = j^{-1}$ . Since the elements of order 4 would have to be mapped to elements of order 4 (since the map is injective), for  $n < 8$ , they would have to be mapped to cycles of length 4. Suppose

$$\begin{aligned} \varphi(i) &= f \\ \varphi(j) &= g \end{aligned}$$

where  $f, g$  are 4-cycles. We would have that

$$\begin{aligned} \varphi(iji^{-1}) &= \varphi(j^{-1}) \\ fgf^{-1} &= g^{-1} \end{aligned}$$

Let  $g = (a \ b \ c \ d)$  be the cycle explicitly. Recall that the conjugation  $fgf^{-1}$  turns the left hand side into

$$(f(a) \ f(b) \ f(c) \ f(d)) = (d \ c \ b \ a)$$

For these two cycles to be equal,  $f$  must map  $a, b, c, d$  again to  $a, b, c, d$  in some way. Since  $k = ij$ , so too would  $k$ . Then  $H$  would be acting only on 4 elements in total, and we would have a faithful map

$$\varphi' : H \rightarrow S_4$$

This is impossible, since the quaternion subgroup is not a subgroup of  $S_4$ . (Suppose it was - it has 6 elements of order 4. These are all the 4-cycles in  $S_4$ . If a subgroup contained them, it would contain their squares as well, and this shows that there would be more than 8 elements in the subgroup.)

Finally, we conclude that  $i, j$  cannot be mapped to 4-cycles. Since  $\varphi(i)$  must have order 4, we conclude that  $H$  cannot faithfully operate on sets of size 4, 5, 6, 7.

Taking  $n = 8$ , we can of course let the group act on itself by left multiplication  $m_g$ . Multiplying by any non-identity element changes at least the identity ( $m_g(1) \neq 1$  if  $g \neq 1$ ). We conclude that  $H$  can faithfully on a set of order minimally 8.

## Result

4 of 4

Note that  $|G|$  must divide  $|S_n|$ .

- a) For  $|D_4| = 8$ , one can take  $n = 4$ . Let it act on the vertices of a square.
- b) For  $|D_6| = 12$ , we see that  $n \geq 4$ .  $D_6$  has an element of order 6, so it can't be  $n = 4, 5$ . Letting it act on a hexagon, shows that we can take  $n = 6$ .
- c) Since the map is faithful, we would need to map  $i, j, k$  to elements of order 4. Up to  $n < 8$  these are the 4-cycles. We show that this is impossible. Letting the group act on itself shows that  $n = 8$  works.

## 8. a

Consider the multiplicative group  $\mathbb{F}_p^\times$  and consider the automorphisms of a cyclic group  $C_p$ . If

$$\theta : C_p \rightarrow C_p$$

is an automorphism, then it preserves orders, so in particular it must map generators to generators. A cyclic group has  $\phi(p) = p - 1 = |\mathbb{F}_p^\times|$  generators (i.e. all non-trivial elements of  $C_p$ ). Let  $g$  be a generator. The automorphisms go like

$$\begin{aligned} g &\mapsto g \\ g &\mapsto 2g \\ &\dots \end{aligned}$$

Thus the automorphisms correspond naturally to left multiplication  $m_a(x) = ax$ . Indeed, define

$$\begin{aligned} \varphi : \mathbb{F}_p^\times &\rightarrow \text{Aut}(C_p) \\ \varphi(a) &= m_a \end{aligned}$$

This is a homomorphism:

$$\varphi(ab) = m_{ab} = m_a m_b = \varphi(a) \varphi(b)$$

It is obviously a bijection between the two sets. We conclude that this is a bijective correspondence between the two groups.

## Result

2 of 2

Note that the automorphisms are just left multiplication by a non trivial element,  $m_a(x) = ax$ . Mapping  $a \in \mathbb{F}_p^\times$  to  $m_a$  gives the desired correspondence.

## 9. a



A group is defined as an algebraic structure that consists of a set of elements equipped with an operation that combines any two elements to form a third element.

[Comment](#)

### Step 2 of 3 ^

Three sheets of rectangular paper  $S_1, S_2, S_3$  are made into stack.

Let ' $G$ ' be a group of all symmetries of this configuration, including symmetries of the individual sheets as well as the permutation of the set of sheets

To determine: the order of ' $G$ ':

Let,  $S_n$  be the group of symmetry for the rectangular paper

Now the order for this symmetric group for which it attains its identity element is given at;

$$n = 1$$

Therefore,  $|G| = 1$

To find; the kernel of the map,

$$G \rightarrow S_3$$

Consider the permutation of the set  $\{S_1, S_2, S_3\}$

Now the kernel for the group is same as that of the order of the group

Hence, the kernel for this set is given as  $\{S_1\}$

## Section 12

1. a

The groups of symmetries of the dodecahedron and the icosahedron are isomorphic because its two solids are dual to each other they have the same symmetry group. The order of the group of direct symmetries of all rotations is,

$$\begin{aligned} |S_d(D)| &= 20 \cdot 3 \\ &= 60 \end{aligned}$$

The elements are 4 rotations by multiples of  $\frac{2\pi}{5}$  about centres of 6 pairs of opposite faces, it has 24 elements.

1 rotation by  $\pi$  about centres of 15 pairs of opposite edges, it has 15 elements.

2 rotations about 10 pairs of opposite vertices, it has 20 elements.

And

There is one identity element.

So, number of element in dodecahedron and the icosahedron is,

$$1 + 24 + 15 + 20 = 60$$

Together with the identity this accounts for all 60 elements.

So,

$$S_d(D) \approx A_5$$



In fact one can embed five cubes in the dodecahedron which are permuted by each rotation.

One may embed five tetrahedral partitioning the 20 vertices and these are permuted also.

Thus,  $S_d(D) \approx A_5$  and  $S(D) \cong A_5 \times \langle J \rangle$

## 2. a

To describe the orbits of poles for the group of rotations of an octahedron;

[Comment](#)

### Step 2 of 4 ^

Octahedron group  $O$ : the group of 24 rotational symmetries of a cube or an octahedron.

There are  $|G| = 24$  rotational symmetries for an octahedron.

That is,

$$\begin{aligned} |G| &= |\text{stabilizer}| |\text{orbit}| \\ &= 24 \end{aligned}$$

This can be split into three orbits for edges, vertices and faces.

Note that  $r_i$  for the size of the stabilization group and  $n_i$  for the size of the orbit.

In particular,

Orbit for edges as shown below;

$$r_i = 2$$

And,

$$n_i = 12$$

Orbit for Vertices as shown below;

$$r_i = 3$$

And,

$$n_i = 8$$

[Comment](#)

### Step 4 of 4 ^

Orbit for faces as shown below;

$$r_i = 4$$

And,

$$n_i = 6$$

Therefore, edge pole orbit order  $\boxed{2}$ , vertex pole orbit order  $\boxed{3}$ , and face pole orbit order  $\boxed{4}$ .

## 3. a

**Given:**  $O$  be the group of rotations of a cube, and let  $S$  be the set of four diagonal lines connecting opposite vertices.

**Solution:** Now by the given condition  $O$  acts on  $S$ .

Note that if  $G$  is a group then there is a bijective correspondence between operations of  $G$  on the set  $S = \{1, 2, \dots, n\}$ .

Therefore there is a homomorphism from  $O$  to  $S_4$ .

Let us consider the homomorphism be

$$f : O \rightarrow S_4.$$

First we show that  $f$  is an injective homomorphism.

Let us assume  $x \in O$  such that  $x$  fixes all four diagonals. Now let us consider two cases below.

**Case-1:** If  $x$  acts trivially on the vertices. Then  $x$  fixes two endpoints of each diagonal or interchanges the two endpoints of each diagonal.

**Case-2:** Now if  $x$  acts non-trivially on the vertices, then we can pick three pairs of opposite vertices such that  $x$  interchanges one of the pairs of vertices, or such that  $x$  interchanges all three pairs of vertices.

In each of the two above cases give  $\mathbb{R}^3$  a basis that points along the three diagonals connecting these pairs of vertices, then note that the matrix for  $x$  in this basis has an  $2n - 1$ , where  $n \in \mathbb{N}$ , number of  $-1$  on the diagonal, and so has determinant  $-1$ .

So this is an impossibility, since  $O$  is a subgroup of  $SO_3(\mathbb{R})$ .

This proves that  $f$  is injective.

Now we show that  $f$  is surjective.

For prove that  $f$  is surjective, look at the faces of the cube. Now notice that  $O$  acts transitively on the set of faces, and each face has stabilizer  $C_4$ .

Hence we have

$$\begin{aligned} |O| &= |O_f| \cdot |O.f| \\ &= 4 \times 6 \\ &= 24, \quad \text{by counting formula.} \end{aligned}$$

Now we know that  $O$  acts like  $S_4$  on the set of diagonals, hence the stabilizer of one of the diagonals is the same as fixing one index in the set of indexes  $\{1, 2, 3, 4\}$  and equal to  $S_3$ .

This proves that  $f$  is surjective. This completes the solution.

## Result

Considering a homomorphism  $f : O \rightarrow S_4$  we show that  $f$  is both injective and surjective.

4. a

To prove that:  $H = T$

Let the group of rotations of the cubes be defined as,

$$G = O$$

$T$  is the tetrahedral group of 12 rotational symmetries of a tetrahedron and  $G = O$  is the octahedral group of 24 rotational symmetries of a cube or an octahedron.

While on the rotation of the two cubes one can form two tetrahedral and one of the tetrahedral is denoted as the subgroup  $H$ .

[Comment](#)

---

Step 2 of 2 ^

Order of group  $G = O$  is 24.

Subgroup of  $G = O$  which is,  $H$  also equals the order of the group.

Order of tetrahedral group is 12.

Since, the tetrahedral is formed from the cube which refers the group  $G$  it must also be a subgroup.

So one can conclude that both the tetrahedral and  $H$  are the subgroup of the group  $G = O$  having the same order which is 12.

Thus,

$$H = T$$

Hence proved

5. a

**To Prove:** The icosahedral group has a subgroup of order 10.

**Proof:** Let us assume that  $H$  be the icosahedral group.

Let us now consider  $a$  be the rotation about a line through the centers of two faces by an angle of  $\frac{2\pi}{5}$ . Let  $b$  be the rotation about a line joining two opposite edges and perpendicular to the above considered line by an angle  $\pi$ .

So we have

$$a^5 = 1 \quad \text{and} \quad b^2 = 1.$$

Now geometrically notice that

$$ba = a^{-1}b.$$

Therefore note that

$$H = \langle a, b \mid a^5 = b^2 = 1, ba = a^{-1}b \rangle.$$

This follows that  $H$  is isomorphic to the Dieder group of order 10.

Hence  $H$  has order 10.

This completes the proof.

---

**Result**

2 of 2

Basically we have shown that Icosahedral Group is isomorphic to the Dieder group of order 10.

6. a

A subset  $H$  of a group  $G$  is a subgroup if;

$$e \in H$$

And, if  $x, y \in H$ , then  $x \times y \in H$

Finally, if  $x \in H$ , then  $x^{-1} \in H$

**a.**

Now, use the Lagrange's theorem which states that;

For any finite group  $G$  the order of every subgroup  $H$  of  $G$  divides the order of  $G$

Here the regular tetrahedron has order 24

Now, further the orders of  $T, D_2, C_3, C_2, C_1$  has an order of 12, 4, 3, 2, 1 respectively. Thus from the above stated Lagrange's theorem it is clear that the order of  $T, D_2, C_3, C_2, C_1$  divides the order of the regular tetrahedron

Therefore, the subgroups are  $T, D_2, C_3, C_2, C_1$  as stated below;

Name	Structure	Order
$T$	$A_4$	12
$D_2$	$D_2$	4
$C_3$	$Z_3$	3
$C_2$	$Z_2$	2
$C_1$	$Z_1$	1

**b.**

Again, use the Lagrange's theorem which states that;

For any finite group  $G$  the order of every subgroup  $H$  of  $G$  divides the order of  $G$

Here the regular icosahedron has order 120

Now, further the orders of  $C_1, C_2, C_3, C_5, D_2, D_3, D_5, I, S_{10}, S_2, S_6, T$  has an order of 1, 2, 3, 5, 4, 6, 10, 60, 10, 2, 6, 12 respectively. Thus from the above stated Lagrange's theorem it is clear that the order of  $C_1, C_2, C_3, C_5, D_2, D_3, D_5, I, S_{10}, S_2, S_6, T$  divides the order of the regular icosahedron

Therefore, the subgroups are  $C_1, C_2, C_3, C_5, D_2, D_3, D_5, I, S_{10}, S_2, S_6, T$  as stated below;

Name	Structure	Order
$C_1$	$Z_1$	1
$C_2$	$Z_2$	2
$C_3$	$Z_3$	3
$C_5$	$Z_5$	5
$D_2$	$D_2$	4
$D_3$	$S_3$	6
$D_5$	$D_5$	10
$I$	$A_5$	60
$S_{10}$	$Z_{10} = Z_2 \times Z_5$	10
$S_2$	$Z_2$	2
$S_6$	$Z_6 = Z_2 \times Z_3$	6
$T$	$A_4$	12

7. a

Icosahedron is defined as those solid figures consisting of twenty plane faces with sides as equilateral triangle.

[Comment](#)

Step 2 of 6 ^

The 12 points  $(\pm 1, \pm \alpha, 0)^T, (0, \pm 1, \pm \alpha)^T, (0, \pm \alpha, \pm 1)^T$  form the vertices of a rectangular icosahedron if  $\alpha > 1$  is chosen suitably.

To verify: the above statement and to determine  $\alpha$

Consider the 12 points from the three categories given above.

Now, using the distance formula, if they are in the same category, then there will be three possibilities for square distances as given below:

Vertex	Change sign of 1	Change sign of $\alpha$	Change sign of both
Square Distance	4	$4\alpha^2$	$4(\alpha^2 + 1)$

If all these are in different categories, then there will be two possibilities for square distances. If the coordinates are  $(x_1, x_2, x_3)$  and  $(y_1, y_2, y_3)$ , then there is only one  $i$  such that  $x_i, y_i \neq 0$  and  $x_i, y_i$  must have different absolute values for each  $i$ , this gives the below table;

Vertex	$x_i, y_i$ have same sign	$x_i, y_i$ have different sign
Square Distance	$2(\alpha^2 - \alpha + 1)$	$2(\alpha^2 + \alpha + 1)$

Here for a given vertexes, there are four vertexes each with square distance of each from above.

Now, find  $\alpha$ .

For this there must be five vertexes of shortest distance from  $v_0$ ; by comparing the possible distances above, this shortest square distance must be equal to both 4 and  $2(\alpha^2 - \alpha + 1)$

This forms the equation;

$$\begin{aligned}
 2(\alpha^2 - \alpha + 1) &= 4 \\
 \alpha^2 - \alpha - 1 &= 0 \\
 \alpha &= \frac{1 + \sqrt{5}}{2}
 \end{aligned}$$

Now, by choosing the root greater than 1, this shows each vertex has exactly five neighbors of distance 2 away from it.

Claim that the polyhedron formed by connecting vertexes of distance 2 away from each other forms an icosahedron.

Now, show that each face formed by the edges is a congruent equilateral triangle. This is true since any face formed by neighboring vertexes, which are distance 2 away from each other by the above.

Each vertex moreover has the same number of faces meeting there since every vertex  $v, v'$  forming an edge and claim that there are only two faces intersecting at that edge.

Now, it remains to show that there are only two vertexes  $w$  that are of distance 2 from both.

Now, suppose;

$$\begin{aligned}
 v_i &= v'_i \\
 &= 0
 \end{aligned}$$

And, further consider subscript mod 3. Then  $v_{i+1} = -v'_{i+1}$  with absolute value 1, so;

$$\begin{aligned}
 w_{i+1} &= 0 \\
 |w_{i+2}| &= 1 \\
 |w_i| &= 3
 \end{aligned}$$

Next;

$$v_{i+1} = -v'_{i+2}$$

Hence,  $w_{i+2}$  must have the same sign as  $v_{i+2}, v'_{i+2}$

Finally,  $w_i$  can have either sign since;

$$\begin{aligned}
 v_i &= v'_i \\
 &= 0
 \end{aligned}$$

Therefore, there are only two vertexes that are of distance 2 from  $v, v'$

8. a



The crystallographic restriction theorem is the basic form which was based on the observation that the rotational symmetries of a crystal are usually limited to 2-fold, 3-fold, 4-fold, and 6-fold.

---

[Comment](#)

---

Step 2 of 2 ^

Consider that the group of displacement in the given plane has more than one centre of rotation, and that the only number of rotations that can occur is 2, 3, 4 and 6.

It can be seen that, the sum of all the interior angles always divided by its number of sides will be divisor of  $360^\circ$ , that is;

$$\frac{180^\circ(x-2)}{x} = \frac{360^\circ}{y}$$

Here,  $y$  is an integer

Thus, symmetry will be possible only if;

$$\frac{2x}{x-2} = y$$

Now, this will hold true for 1-fold, 2-fold, 3-fold, 4-fold, and 6-fold symmetry. This means that the restriction does not occur for  $x > 6$

**Therefore, the crystallographic restriction for three-dimensional crystallographic groups that is, a rotational symmetry of a crystal has order 2, 3, 4, or 6**

## Miscellaneous Problem

1. a

Let the lattice be denoted by  $L$

Consider  $G$  to be a group of symmetries.

The translation subgroup of  $G$  is the lattice on the basis vectors  $(1,0)$  and  $(0,1)$ .

Then there are three possibilities;

First is that either  $L$  is trivial, that is,  $G$  is rigid motion

Second is that  $L$  is generated by just one translation and the other one can be a glide

And, the third one is that  $L$  is generated by two linearly independent translations.

---

[Comment](#)

---

Step 2 of 2 ^

Further let  $G$  be a discrete group of rigid motions of the group of rigid motions of the plane.

Then the point  $\bar{G}$  of  $G$  will be fixed

Consider  $G$  be a two-dimensional crystallographic group.

So, clearly the point group  $\bar{G}$  is a subgroup of symmetries of a two dimensional lattice.

Since,  $G$  is a two-dimensional

So, the translations will be of the form of  $C_n$  or  $D_n$

**Thus, there will be two translations.**

2. a

- [a]) Let  $\text{Aut } G$  be the set of automorphisms of a group  $G$ . We prove that it is a group.

Let  $\varphi, \psi$  be two automorphisms. Then

$$\begin{aligned}(\varphi \circ \psi)(ab) &= \varphi(\psi(ab)) \\&= \varphi(\psi(a)\psi(b)) \\&= \varphi(\psi(a))\varphi(\psi(b)) \\&= (\varphi \circ \psi)(a)(\varphi \circ \psi)(b)\end{aligned}$$

Thus the composition of homomorphisms is a homomorphism.  $\varphi, \psi$  are bijections from  $G$  onto  $G$ . Since bijections compose into bijections, we conclude that we have a bijective homomorphism between  $G$  and  $G$ . Thus the set  $\text{Aut } G$  is closed on composition.

The associative law follows from the associativity of composition. The identity element is the identity function. Finally, note that every element has an inverse. Indeed, since automorphisms are bijections, each one has an inverse. This proves all the group axioms, and so the set is a group.

- [b]) Consider the map  $\varphi : G \rightarrow \text{Aut } G$  defined by

$$g \mapsto \{\text{conjugation by } g\}$$

We show that this is a homomorphism. Denote by  $c_g(x) = gxg^{-1}$ , the map which conjugates by  $g$ . Let  $g, h \in G$ . Then

$$\begin{aligned}\varphi(gh)(x) &= c_{gh}(x) \\&= (gh)x(gh)^{-1} \\&= gc_h(x)g^{-1} \\&= c_g(c_h(x)) \\&= (\varphi(g) \circ \varphi(h))(x)\end{aligned}$$

This shows that the map is a homomorphism.

Its kernel is the set of  $g$ , such that

$$\varphi(g)(x) = c_g(x) = x$$

So, conjugation by  $g$  should leave every element  $x \in G$  fixed. This means that

$$\begin{aligned}gxg^{-1} &= x \\gx &= xg, \quad \text{for all } x \in G\end{aligned}$$

We conclude that  $g \in Z(G)$  and that  $\ker \varphi = Z(G)$ , the center of the group.

- [c]) We prove that the group of *inner automorphisms* of the type in part b) is a normal subgroup of  $\text{Aut } G$ . Let  $c_g$  be an inner automorphism and let  $\varphi$  be an arbitrary automorphism. We want to show that

$$\varphi \circ c_g \circ \varphi^{-1}$$

is again an inner automorphism (i.e. conjugation by some element). Let's check how it maps. We find that

$$\begin{aligned}(\varphi \circ c_g \circ \varphi^{-1})(x) &= (\varphi \circ c_g)(\varphi^{-1}(x)) \\&= \varphi(g\varphi^{-1}(x)g^{-1}) \\&= \varphi(g)x\varphi(g)^{-1} \\&= c_{\varphi(g)}(x)\end{aligned}$$

We find that indeed, for arbitrary  $\varphi$ ,  $\varphi c_g \varphi^{-1}$  is again an inner automorphism. We conclude that the subgroup of inner automorphisms is normal.

## Result

a) One checks the groups laws.

b) Denote by  $c_g(x) = gxg^{-1}$ . One proves that  $\varphi(gh)(x) = (\varphi(g) \circ \varphi(h))(x)$ .

Its kernel is the set of  $g$  such that  $c_g(x) = gxg^{-1} = x$  and this means that  $g \in Z(G)$ .

c) Note that  $\varphi \circ c_g \circ \varphi^{-1} = c_{\varphi(g)}$ . This shows that the subgroup is normal.

### 3. a

- [a)] Consider the cyclic group  $C_4$ . Any automorphism preserves the orders of elements, and so must map a generator into a generator. The value of the automorphism on the generator also completely determines it. We have two generators, and so two possible automorphisms. Therefore

$$\text{Aut } G \cong C_2$$

Since the group is abelian, any inner automorphism is trivial ( $gxg^{-1} = x$  always). Thus

$$\text{Inn } G \cong \{1\}$$

## Step 2

2 of 6

- [b)] Consider the cyclic group  $C_6$ . Automorphisms map generators into generators. We have  $\phi(6) = 2$  generators, and so two possible automorphisms. Therefore

$$\text{Aut } G \cong C_2$$

Since the group is abelian, any inner automorphism is trivial. Thus

$$\text{Inn } G \cong \{1\}$$

- [c)] Consider the group  $C_2 \times C_2 = \{1, a, b, c\}$ .

This group has 2 generators (any two of the three non-trivial elements). There are  $3 \times 2 = 6$  ways to choose how to map the generators, and are all valid:

$$\begin{aligned}\varphi_1(a) &= a, & \varphi_1(b) &= b \\ \varphi_2(a) &= b, & \varphi_2(b) &= a \\ \varphi_3(a) &= c, & \varphi_3(b) &= b \\ \varphi_4(a) &= b, & \varphi_4(b) &= c \\ \varphi_5(a) &= a, & \varphi_5(b) &= c \\ \varphi_6(a) &= c, & \varphi_6(b) &= a\end{aligned}$$

Note that the  $\varphi_i$  just permute the elements  $a, b, c$ . We conclude that  $\text{Aut } G \cong S_3$ .

Since the group is abelian, any inner automorphism is trivial. Thus

$$\text{Inn } G \cong \{1\}$$

- [d)] Consider the group  $D_4$  generated by rotation  $x$  and reflection  $y$ .

We know that it is generated by any of the two rotations  $x, x^3$  and any of the four reflections  $y, yx, yx^2, yx^3$ .

An automorphism will necessarily map  $x$  to either  $x, x^3$  and a reflection will be mapped to a reflection as well.

This means that there are 2 choices where to map  $x$  and 4 choices where to map  $y$  for a total of  $2 \times 4 = 8$  choices.

Let us find which group this is. Choose the two automorphisms:

$$\begin{aligned}\varphi(x) &= x^3, & \varphi(y) &= y \\ \theta(x) &= x, & \theta(y) &= yx\end{aligned}$$

Then it is easily seen that  $\varphi^2 = \theta^4 = \text{id}$ , the identity. Also, we find that

$$\begin{aligned}(\varphi \circ \theta)(x) &= x^3 = (\theta^{-1}\varphi)(x) \\ (\varphi \circ \theta)(y) &= yx^3 = (\theta^{-1}\varphi)(y)\end{aligned}$$

Since this holds for the generators, it holds for all other elements as well. We see that the automorphism group is of order 8 and generated in the same way as the dihedral group. We conclude that

$$\text{Aut } D_4 \cong D_4$$

Now consider the inner automorphisms. We have seen in the previous exercise that they make a normal subgroup of  $\text{Aut } G$ , and in fact

$$\text{Inn } G \cong \frac{G}{Z(G)}$$

for any group  $G$ . Taking  $G = D_4$ , we know its center is  $Z(D_4) = \{1, x^2\}$ , and so the quotient group is isomorphic to

$$\text{Inn } G \cong \{\bar{1}, \bar{x}, \bar{y}, \bar{xy}\}$$

where the bars denote that these are cosets. We see that this is not a cyclic group, and so must be  $C_2 \times C_2$ .

- [e]) Consider the quaternion group  $H$ . It has six elements of order 4, and any pair of them will serve as generators. Thus it is enough to define an automorphism on them. The first generator  $g_1$  can be mapped to any of the 6 elements, whereas the second one cannot be mapped into  $\pm g_1$ . This gives

$$6 \times 4 = 24$$

possibilities. As we have seen before, 24 is the order of the symmetric group. We also know that this appears e.g. in tetrahedral symmetry. Imagine that the elements of  $Q_4$  of order 4, namely

$$i, -i, j, -j, k, -k$$

represent opposing faces of the cube. It can be checked that a symmetry of the cube is uniquely defined by defining which two faces it switches. Therefore, every automorphism of  $Q_4$  induces a unique symmetry of the cube.

Since the symmetry group of a cube is  $S_4$  of order 24, we conclude that the automorphism group must be

$$\text{Aut } H = S_4$$

The inner automorphism group is derived as in the previous example. The center of the group  $H$  is

$$Z(H) = \{\pm 1\}$$

Then

$$\begin{aligned} \text{Inn } H &\cong \frac{H}{Z(H)} = \frac{H}{\{\pm 1\}} \\ &\cong \{1, i, j, k\} \end{aligned}$$

which is not cyclic, and therefore isomorphic to  $C_2 \times C_2$ .

## Result

6 of 6

a), b) both have automorphism group  $C_2$  and trivial inner automorphisms.

c) has automorphism group  $S_3$  and trivial inner automorphisms since it is abelian.

d)  $D_4$  has as automorphism group a group isomorphic to  $D_4$ . This is shown by proving the same relations in  $\text{Aut } D_4$  as in  $D_4$ . The inner automorphism is  $G/Z(G) \cong C_2 \times C_2$ .

e) One can assign the elements of order 4,  $i, -i, j, -j, k, -k$  to faces of the cube. Each automorphism corresponds to a unique symmetry. Therefore,  $\text{Aut } H \cong S_4$ , the symmetry of a cube. Inner automorphisms are  $C_2 \times C_2$ .

4. a



(a)

To determine the order of the subgroup  $G_n$ ,

Suppose  $G_n$  be the subgroup of the orthogonal group  $O_n$  of elements that sends the hypercube to itself,

Now, file a claim that, order of the group  $G_n$  is  $2^n n!$ .

That is,

$$|G| = 2^n n!$$

Use the induction to show that  $|G| = 2^n n!$ .

Step-1, the base case, put  $n = 1$

Then,  $x \in \mathbb{R}$  and

$$-1 \leq x \leq +1$$

Then, 1-dimensional hypercube  $C_1$  is the line segment.

Then, the subgroup  $G_1$  of the orthogonal group  $O_1$  contains symmetries of a line segment that are reflection and identity.

Then,

$$\begin{aligned} |G| &= 2^1 \cdot 1! \\ &= 2 \end{aligned}$$

Step-2, the inductive case

Suppose  $k$  be some unspecified number such that  $k \geq 1$  the order  $2^k k!$  of the subgroup  $G_k$  be true.

Then,

$$|G_k| = 2^k k!$$

Then, the  $n$ -dimensional hypercube  $C_n$  has  $2n$  faces and order of  $O_n$  is  $2n$ .

Now, show that the final step  $k+1$  is true.

Then,

$$\begin{aligned} |G_{k+1}| &= 2^{k+1} (k+1)! \\ &= 2 \cdot 2^k (k+1) \cdot k! \\ &= (2^k k!) [2(k+1)] \\ &= |G_k| [2(k+1)] \\ &= |G_k| |G_1| (k+1) \end{aligned}$$

Since,  $|G_k|$  and  $|G_1|$  are true.

Then,

$$\begin{aligned} |G_{k+1}| &= 2^{k+1} (k+1)! \\ &= 2 \cdot 2^k (k+1) \cdot k! \\ &= (2^k k!) [2(k+1)] \\ &= |G_k| [2(k+1)] \\ &= |G_k| |G_1| (k+1) \end{aligned}$$

Hence, it concludes that  $|G_n| = 2^n n!$ .



(b)

To describe the group  $G_n$  explicitly and identify the stabilizer of the vertex  $(1, \dots, 1)$ ,

Suppose  $G_n$  be the subgroup of the orthogonal group  $O_n$  of elements that sends the hypercube to itself and  $v$  be the stabilizer of the vertex.

Then,

$$v = \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{bmatrix}$$

Then, the  $n$  neighboring vertices carried by symmetries fixing  $v$  into each other.

That is,  $n$  neighboring vertices are must permuted by symmetry.

Then, to determine the symmetry completely will be sufficient.

Therefore, the stabilizer of the vertex  $v$  will be isomorphic to a subgroup of the group of permutation  $S_n$ .

Since, the  $n$ -dimensional hypercube  $C_n$  has  $2^n$  vertices.

Then, there will be symmetry can take  $v$  to any one from  $2^n$  vertices of  $n$ -dimensional hypercube  $C_n$ .

Since,  $O_n$  is the orthogonal group elements that sends the hypercube to itself.

Then,

$$|O_v| = 2^n$$

And the order of the group  $G_n$  is,

$$|G_n| = 2^n n!$$

Now, by counting formula

$$\begin{aligned} |G_v| &= \frac{|G_n|}{|O_v|} \\ &= \frac{2^n n!}{2^n} \\ &= n! \end{aligned}$$

Since, the group of stabilizer of the vertex  $v$  is  $|G_v| = n!$  and it is isomorphic to a subgroup of the group of permutation  $S_n$ .

This implies that,

$$G_v \cong S_n$$

Since,

$$|G_n| = 2^n n!$$

Then,

$$\begin{aligned} |G_2| &= 2^2 \cdot 2! \\ &= 4 \cdot 2 \\ &= 8 \end{aligned}$$

And

$$\begin{aligned} |D_4| &= 2 \cdot 4 \\ &= 8 \end{aligned}$$

Then,

$$G_2 \cong D_4$$

Hence, it concludes that the group of stabilizer  $G_v$  of the vertex  $v$  and the group of permutation  $S_n$  are isomorphic.

5. a

!!!

6. a

!!!

7. a

- a) Let  $G = D_3$  be the dihedral group of order 6 and consider its action on the vertices of a triangle. We make a "truth" table whenever  $gs = s$  for  $g \in G$  and  $s \in S$ , the set of vertices of the triangle:

	$v_1$	$v_2$	$v_3$
1	+	+	+
$x$	-	-	-
$x^2$	-	-	-
$y$	+	-	-
$yx$	-	+	-
$yx^2$	-	-	+

- [b] The idea of the above exercise is that we can count the elements by rows, but we may also count them by columns. Let  $G_s$  be the stabilizer of  $s$  and let  $S^g = \{s \in S \mid gs = s\}$ . Thus  $G_s$  corresponds to the rows above, whereas  $S^g$  corresponds to the columns. Thus by a change of summation, we get the desired formula as follows:

$$\begin{aligned}
 \sum_{s \in S} |G_s| &= \sum_{s \in S} \sum_{\substack{g \in G \\ gs=s}} 1 \\
 &= \sum_{\substack{s \in S, g \in G \\ gs=s}} 1 \\
 &= \sum_{g \in G} \sum_{\substack{s \in S \\ gs=s}} 1 \\
 &= \sum_{g \in G} |S^g|
 \end{aligned}$$

- [c] From the previous, we know that

$$\sum_{g \in G} |S^g| = \sum_{s \in S} |G_s|$$

Now, recall the counting formula that states  $|G| = |O_s| |G_s|$  where  $O_s$  is the orbit of an elements  $s$ . Plug this in to get

$$\begin{aligned}
 \sum_{s \in S} |G_s| &= \sum_{s \in S} \frac{|G|}{|O_s|} \\
 &= |G| \sum_{s \in S} \frac{1}{|O_s|}
 \end{aligned}$$

Every orbit has exactly  $|O_s|$  elements in it. Since the orbits are disjoint, every fraction  $\frac{1}{|O_s|}$  in the above sum appears exactly  $|O_s|$  times. We conclude that every orbit contributes to the sum with exactly 1 and so the sum on the right counts the **number of orbits**. We have proven that

$$\sum_{g \in G} |S^g| = |G| (\text{number of orbits})$$

## Result

4 c

a) The truth table is straightforwardly checked.

b) The point which part a) illustrates is that we can change the order of summation. Doing this, we get the statement.

c) Recall that  $|G| = |G_s| |O_s|$  and use part b).

8. a

There are  $70 = \binom{8}{4}$  ways to color the edges of an octagon, with four black and four white. Suppose the group  $D_8$  operates on this set of 70 colorings, and the orbits represent equivalent colorings. Burnside's formula from the previous exercise states that

$$\text{number of orbits} = \frac{1}{|G|} \sum_{g \in G} |S^g|$$

where  $S^g$  is the set of elements  $s$  fixed by  $g$ .

We count the sets by possible actions of the group.  $D_8$  consists of:

- the identity map. Here  $|S^1| = 70$  since every element is fixed by it.
- rotation by  $180^\circ$ . Assign letters B,W to the octagon to signify if it's black or white. This will give a cyclical pattern of letters  $ABCDEFGH$ . After rotating by  $180^\circ$ , it will read

$$EFGHABCD$$

Thus rotation  $180^\circ$  fixes those colorings of the form  $ABCDABCD$ , i.e. when the coloring is repeated twice. We can choose  $A, B, C, D$  in any way, so since there are two colors, there are  $\binom{4}{2} = 6$  elements that rotation by  $180^\circ$  fixes.

- rotation by  $\pm 90^\circ$  fixes colorings  $ABCDEFGH$  if after rotating by  $90^\circ$  we have

$$ABCDEFGH = CDEFGHAB$$

(or vice-versa for  $270^\circ$ ).

We see that only two letters can be chosen arbitrarily. Thus for rotations by  $\pm 90^\circ$ , we have only two choices.

- rotations by  $45^\circ$  which fix no coloring.
- Suppose we have the coloring  $ABCDEFGH$  and consider a reflection through opposing sides. Say the axis of reflection goes through  $A$  and  $E$ . This reflects the coloring into:

$$AHGFEDCB$$

We see that it is necessary and sufficient to have  $B = H, C = G, D = F$  so the coloring becomes of the form

$$ABCDED CB$$

Since  $B, C, D$  come in pairs, we must have  $A = E$ . Finally, these  $A, B, C, D$  can be chosen to arbitrarily, so we have  $\binom{4}{2} = 6$ .

- Finally, consider reflections around vertices. Like above, we find that  $ABCDEFGH$ , gets mapped to

$$HGFEDCBA$$

Thus it is necessary and sufficient that the coloring has the form ABCDDCBA. There are  $\binom{4}{2} = 6$  such colorings.

Recall there is only one identity, one rotation by  $180^\circ$ , two rotations by  $90^\circ$ , four reflections around axes connecting vertices, four reflections around axes connecting opposing edges. Utilizing Burnside's formula, we get:

$$\begin{aligned}\text{number of orbits} &= \frac{1}{|G|} \sum_{g \in G} |S^g| \\ &= \frac{1}{16} (70 + 6 + 2 \cdot 2 + 4 \cdot 6 + 4 \cdot 6) \\ &= \frac{128}{16} \\ &= 8\end{aligned}$$

## Result

4 of 4

We use Burnside's formula. To do so, we must count for each element of  $G = D_8$ , the fixed points (here, the colorings which stay fixed). This is made easier by fixing a certain edge, and then coloring the sides  $ABCDEFGH$ . There are always two colors to choose from. Using combinatorics, we arrive at all the  $|S^g|$  and compute the number of orbits.

## Chapter 7

### Section 1

1. a

The given rule  $g * x = xg^{-1}$  defines operation of  $G$  on  $G$ . The operation satisfies two properties.

1. For  $1, x \in G$ ,

$$1 * x = x1^{-1} = x1 = x$$

2. *Associative law* : For  $g_1, g_2 \in G$

$$g_1 * (g_2 * x) = g_1 * (xg_2^{-1}) = xg_2^{-1}g_1^{-1} = x(g_1g_2)^{-1} = (g_1g_2) * x$$

By 6.7, it defines operation of  $G$  on  $G$ .

#### Result

Show that  $1 * x = x$  and associative law holds under given operation.

2. a

Given  $H$  is subgroup of  $G$  and  $H$  operate on  $G$  by left multiplication i.e. for  $h \in H, s \in G$ ,

$$h * s = hs$$

The orbit of  $s$  consists of

$$O_s = \{hs : h \in H\} = Hs$$

which is a right coset of  $H$  in  $G$ . Thus the orbits for the operation of  $H$  on  $G$  by left multiplication consists of right cosets of  $H$  in  $G$ .

#### Result

2 of 2

The orbits consists of right cosets of  $H$  in  $G$ .

### Section 2

1. a

(a) Let  $a, b, c, d \in \mathbb{F}_3$ , such that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Simplifying gives

$$\begin{bmatrix} a & a+b \\ c & c+d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ c & d \end{bmatrix}$$

This gives  $a = a + c, a + b = b + d, c + d = d$  i.e.  $c = 0, a = d$ . Since

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL_n(\mathbb{F}_3)$$

, the elements which commute with given matrix are

$$Z\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = \left\{ \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} : a \in \{1, 2\}, b \in \{0, 1, 2\} \right\}$$

The centralizer has 6 elements. There are  $\prod_{i=0}^{n-1} (3^2 - 3^i) = 48$  elements in general linear group  $GL_n(\mathbb{F}_3)$  over finite field  $\mathbb{F}_3$ . By counting formula the order of conjugacy class is

$$\frac{48}{6} = 8$$

(b) Let  $a, b, c, d \in \mathbb{F}_5$ , such that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Simplifying gives

$$\begin{bmatrix} a & 2b \\ c & 2d \end{bmatrix} = \begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix}$$

This gives  $2c = c, 2b = b$ . However for  $\mathbb{F}_3$ , this means  $b = c = 0$ . Thus, the elements that commute with given element is

$$Z\left(\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right) = \left\{ \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} : a, d \in \{1, 2, 3, 4\} \right\}$$

i.e. there are  $4^2 = 16$  elements in centralizer. Now there are

$$\prod_{k=0}^{2-1} (5^2 - 5^k) = 480$$

in  $GL_n(\mathbb{F}_5)$ . The order of conjugacy class of given element must be

$$\frac{480}{16} = 30$$

## Result

(a) The order of conjugacy class is 8 (b) The order of conjugacy class is 30.

2. a



Given  $G$  has order of  $21 = 3 \cdot 7$ . The order of element  $1 \neq x \in G$  can be 3 or 7. Now, it is given that it's conjugacy class  $C(x)$  is of order 3. Since

$$|Z(x)| \cdot |C(x)| = |G| \implies |Z(x)| = 7$$

Now it can be shown that  $Z(x)$  is subgroup of  $G$ . Also  $x \in Z(x)$  and the order of  $x$  is the order of cyclic subgroup  $\langle x \rangle$  generated by  $x$ . Order of  $\langle x \rangle$  divides the order of  $Z(x)$ . Since 7 is prime,  $\langle x \rangle$  must be of same order i.e.  $x$  is order 7.

### Result

2 of 2

$|\langle x \rangle| \mid |Z(x)|$  so show that  $Z(x)$  is of order 7. This gives  $x$  of order 7.

### 3. a

Given that group  $G$  of order 12 contains a conjugacy class of order 4. Let  $x$  be an element such that  $|C(x)| = 4$  then

$$|Z(x)| = |G|/|C(x)| = 3$$

This obviously contains 1 and  $x$ . Suppose if  $g \in G$  such that  $gxg^{-1} = x$  then

$$g^{-1}gxg^{-1}g = g^{-1}xg = x$$

which shows that  $g^{-1} \in G$ . Therefore the elements of

$$Z(x) = \{1, x, x^{-1}\}$$

Now the center  $Z(G)$  must commute with every element of  $G$  so

$$Z(G) = \bigcap_{x \in G} Z(x)$$

It is known that there are elements in  $G$  which does not commute with  $x$  (which implies it does not commute with  $x^{-1}$  either). Therefore center must be trivial.

### Result

2 of 2

Show that  $Z(x) = \{1, x, x^{-1}\}$  and  $x \notin Z(G)$ .

### 4. a

Let  $C_a = \{g^{-1}ag : g \in G\}$  be the conjugacy class of  $a$  in  $G$ . Then for given map  $\varphi(x) = x^n$ ,

$$\varphi(g^{-1}ag) = (g^{-1}ag)^n = (g^{-1}ag)(g^{-1}ag) \dots (g^{-1}ag) = g^{-1}a^n g$$

This shows  $\varphi : C_a \rightarrow C_{a^n}$ .

### Result

$\varphi : C_a \rightarrow C_{a^n}$  where  $\varphi$  is surjective.

Method 2.

Consider  $G$  be a group and  $\phi$  be the  $n$ th power map:  $\phi(x) = x^n$ .

[Comment](#)

## Step 2 of 2 ^

Assume  $x, y \in G$  such that  $x$  and  $y$  are conjugate,

Then for some  $g \in G$  such that,

$$x = gyg^{-1},$$

And, let  $e$  be the identity element.

Then,

$$\begin{aligned} x^n &= (gyg^{-1})(gyg^{-1})\dots(gyg^{-1})(gyg^{-1}) \\ &= gy(g^{-1}g)y(g^{-1}g)yg^{-1}\dots gy(g^{-1}g)yg^{-1} \\ &= gy(e)y(e)yg^{-1}\dots gy(e)yg^{-1} \\ &= gy^ng^{-1} \end{aligned}$$

Therefore, if  $x$  and  $y$  are conjugate, their powers  $x^n$  and  $y^n$  are also conjugate.

Hence,  **$n$ th power map  $\phi(x) = x^n$  gives a map on conjugacy classes.**

## 5. a

Given  $G$  is a group of matrices of form

$$\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}$$

where  $x > 0$ . Let

$$A = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$$

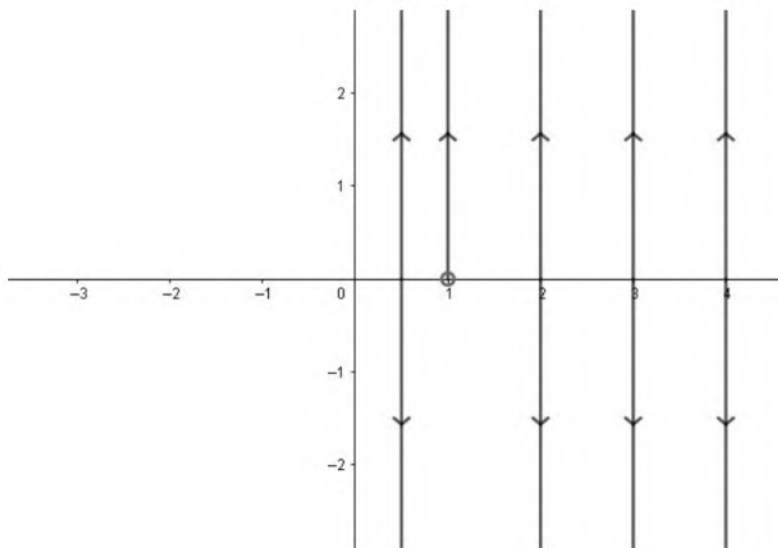
where  $a > 0$ . Then

$$\begin{aligned} C_A &= \left\{ \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} : \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \in G \right\} \\ &= \left\{ \begin{bmatrix} a & \frac{b+ay-y}{x} \\ 0 & 1 \end{bmatrix} : x > 0, y \in \mathbb{R} \right\} \end{aligned}$$

Now, let  $c \in \mathbb{R}$ , for fixed  $x$ , choose  $y$  as

$$b + y(a - 1) = cx \implies y = \frac{cx - b}{a - 1}$$

If  $a \neq 1$  then there exists  $y$  for any value of  $c$ . Now if  $a = 1$  then  $c$  attains any value in  $(0, \infty)$  or  $(-\infty, 0)$  depending on the sign of  $b$ .



## Result

2 of 2

The conjugacy class can be visualized as straight vertical line on positive  $x$  axis. At  $x = 1$ , the conjugacy class is half vertical line without touching  $(1, 0)$ .

6. a

Consider the conjugacy classes in the group  $M$  of isometries of the plane.

[Comment](#)

Step 2 of 2 ^

Let  $u_\theta$  be the reflection across the line passing through origin and making an angle  $\theta$  with the positive  $x$ -axis.

And,

The reflection  $u_\theta$  consists of rotation of the plane to shift the line of reflection onto the  $x$ -axis, reflection across the  $x$ -axis and reverse rotation of the plane to its previous position.

Thus,

$$\begin{aligned} u_\theta &= r_\theta u r_\theta^{-1} \\ &= r_\theta u r_{-\theta} \end{aligned}$$

Therefore, **the conjugacy classes are the reflections across any line in the group  $M$  of isometries of the plane.**

7. a

Given, a group of order 10. The conjugacy class of  $a \in G$  is given by

$$C_a = \{b \in G : g^{-1}ag = b \text{ for some } g \in G\}$$

$$1 + 1 + 1 + 2 + 5 :$$

Now there are three elements for which order of conjugacy class is 1. Since  $a \in C_a$  so  $g^{-1}ag = a$  for all  $g \in G$ . This shows there three elements which commute with every element so  $|Z| = 3$ . Now since  $Z$  is normal subgroup of  $G$ , by Cayley theorem  $|Z| \mid |G|$  however  $3 \nmid 10$  hence this cannot be the class equation

## Step 2

2 of 4

$$1 + 2 + 3 + 4 :$$

Here, for some  $a \in G$ ,  $|C_a| = 3$  but since  $|C_a||Z(a)| = |G|$  however  $3 \nmid 10$ . Thus this cannot be class equation.

$$1 + 1 + 2 + 2 + 2 + 2 :$$

Here again, we note that  $|Z| = 2$  and since  $Z$  is normal subgroup, the quotient group  $G/Z$  has order  $|G/Z| = |G|/|Z| = 5$  which is cyclic group (being of order prime). It can be shown that if  $G/N$  is cyclic then  $G$  must be Abelian. Abelian group has only class equation  $1 + 1 + \dots + 1$  since every element commutes with each other. Therefore this cannot be the class equation.

## Result

4 of 4

$1 + 2 + 2 + 5$  is the required class equation of group 10. Dihedral group  $D_5$  of order 10 has this class equation precisely.

## 8. a

**8 :** Now, for group  $G$  of order 8, the center  $Z$  cannot be of order 4 since  $G/Z$  would be of order  $8/4 = 2$  which would imply  $G/Z$  cyclic which in turn implies  $G$  is abelian.

Now, suppose if  $Z$  were of order 1. Since  $|C_a|$  must divide  $|G|$  for all  $a \in G$ ,  $C_a$  must be of even order. So  $1 + \text{even} + \dots + \text{even} \neq 10$ . So  $Z$  cannot be of order 1.

Another number dividing 8 is 2 so  $Z$  must be of order 2. Now if  $x \notin Z$ , then  $Z(x)$  contains  $Z$  as well as  $x$  so  $Z(x)$  must be of order 4 since  $|Z(x)|$  divides 8 the order of group. Thus  $C_x$  must be of order 2. So the class equation is of

$$2 + 2 + 2 + 2$$

**21 :** For group  $G$  of order  $21 = 3 \times 7$ , the center  $Z$  cannot be of order 3 or 7 since  $|G/Z|$  would be prime which would mean that it is cyclic hence  $G$  abelian. Therefore, the center of the group must be trivial.

Since  $1 + 3n \neq 21$  and  $1 + 7n \neq 21$  for any positive integer  $n$ , the conjugacy classes must be of both order 7 and 3. Now,

$$1 + 3n + 7m = 21$$

has only one solution  $m = n = 2$  i.e. class equation is

$$1 + 3 + 3 + 7 + 7$$

## Result

3 of 3

Class equation for group of order 8 is  $2 + 2 + 2 + 2 = 8$ . The class equation for group of order 21 is  $21 = 1 + 3 + 3 + 7 + 7$

9. a

(a) The quaternion group  $H$  and (b) dihedral group  $D_4$  are non abelian groups of order 8. The quaternion group has center

$$Z(H) = \{1, -1, \pm i, \pm j, \pm k\}$$

and the dihedral group  $D_4$  has center

$$Z(D_4) = \{(1), (1\ 3)(2\ 4)\}$$

Each centralizer of element contains center along with element so order of centralizer must be 2. The class equation must be

$$2 + 2 + 2 + 2 = |G| = 8$$

(Also see problem 2.9)

(c)  $D_5$  is dihedral group of order 10. The order of center can be 2 or 5. This implies  $D_5/Z(D_5)$  is of order  $10/5 = 2$  or  $10/2 = 5$  which must be cyclic which implies  $D_5$  is abelian (contradiction). So the order of  $Z(D_5)$  cannot be 3 or 5 so it must be 1. Now,

$$1 + 2n + 5m = 10$$

has only one solution i.e.  $n = 2, m = 1$  so the class equation is

$$1 + 2 + 2 + 5 = 10$$

(Also see problem 2.6)

(d) Now subgroup of  $GL_2(\mathbb{F}_3)$  which consists of upper triangular matrices consists of 12 elements. Since divisors of 12 are 1, 2, 3, 4 and 6. Since  $12/6 = 2$  and  $12/4 = 3$  which are prime, the order of center cannot be 6 and 4.

The only matrices that commute with every other matrix is

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

therefore the size center is 2. The conjugacy classes are listed as below.

$$\begin{aligned} C\left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right) &= \left\{ \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\} \\ C\left(\begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}\right) &= \left\{ \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \right\} \\ C\left(\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\right) &= \left\{ \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} \right\} \\ C\left(\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\right) &= \left\{ \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \right\} \end{aligned}$$

Therefore the class equation is

$$1 + 1 + 2 + 2 + 3 + 3 = 12$$

## Result

(a) and (b)  $2 + 2 + 2 + 2$  (c)  $1 + 2 + 2 + 5$  (d)  $1 + 1 + 2 + 2 + 3 + 3 = 12$

10. a

(a)

Let  $A$  be an element of  $SO_3$  that represents a rotation with angle  $\pi$ .

To describe the centralizer of  $A$  geometrically;

Suppose that  $A$  centralizes  $SO_3$  and it is known that element of this group belongs to  $M_3(\mathbb{C})$ .

Then,  $A$  must be commutes every  $R$ , where  $R$  is a rotation matrix.

$$AR = RA$$

Rotation  $R$  can be defined as,

$$R = P \left( \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \oplus I_1 \right) P^T$$

Where,  $P$  is permutation matrix.

Let  $P = I$  and partition  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .

Where,  $a$  is  $2 \times 2$  matrix.

$$R = \begin{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & 0 \\ 0 & 1 \end{pmatrix}$$

Suppose  $C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$  then this can be written as,

$$R = \begin{pmatrix} C & 0 \\ 0 & 1 \end{pmatrix}$$



Now substitute  $A$  and  $R$  in  $AR = RA$ , to get

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

This implies that,

$$Cb = b \text{ and } Cc = c$$

This implies that,

$$b = 0, c = 0$$

$$\text{Thus, } A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$$

Therefore, centralize  $A$  must be diagonal matrix and  $A$  is a scalar multiple of  $I_3$ .

(b)

To determine the centralizer of the reflection  $r$  about the  $e_1$ -axis in the group  $M$  of isometries of the plane;

Reflection  $r$  about the  $e_1$ -axis in the group  $M$  of isometries of the plane can be determine by,

$$r(x) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Let  $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  and  $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  belongs to the centralizer of  $A$  in group  $M$ .

Then,

$$AX = XA$$

Substitute matrix  $A$  and  $X$  in the above,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ -c & -d \end{pmatrix} = \begin{pmatrix} a & -b \\ c & -d \end{pmatrix}$$

Now equating corresponding entries shows that  $a = a, b = -b, -c = c, -d = -d$ .

This implies that  $b = 0, c = 0$ ,  $a$  and  $d$  can be any number.

Hence, the centralizer of the reflection  $r$  about the  $e_1$ -axis in the group  $M$  of isometries of the

plane is  $\left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} : a, d \in \mathbb{R} \right\}$ .

11. a

Given the elements of  $GL_3(\mathbb{R})$ , the centralizer are those elements which commute with the given element.

(a) Let

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in GL_3(\mathbb{R})$$

such that

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

This gives us

$$\begin{bmatrix} a & 2b & 3c \\ d & 2e & 3f \\ g & 2h & 3i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 2d & 2e & 2f \\ 3g & 3h & 3i \end{bmatrix}$$

Comparing each element, we get  $b = c = d = f = g = h = 0$ . Therefore the centralizer is

$$\left\{ \begin{bmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{bmatrix} : a \neq 0, e \neq 0, i \neq 0 \right\}$$

(b) Let

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in GL_3(\mathbb{R})$$

such that

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

This gives us

$$\begin{bmatrix} a & b & 2c \\ d & e & 2f \\ g & h & 2i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 2g & 2h & 2i \end{bmatrix}$$

Comparing each element, we get  $c = f = g = h = 0$ . Therefore the centralizer is

$$\left\{ \begin{bmatrix} a & b & 0 \\ d & e & 0 \\ 0 & 0 & i \end{bmatrix} : i \neq 0, ae - bd \neq 0 \right\}$$

(c) Let

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in GL_3(\mathbb{R})$$

such that

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

This gives us

$$\begin{bmatrix} a & a+b & c \\ d & d+e & f \\ g & g+h & i \end{bmatrix} = \begin{bmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{bmatrix}$$

Comparing each element, we get  $d = f = g = 0$  and  $e = q$ . Therefore the centralizer is

$$\left\{ \begin{bmatrix} a & b & c \\ 0 & a & 0 \\ 0 & h & i \end{bmatrix} : a \neq 0, i \neq 0 \right\}$$

(d) Let

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in GL_3(\mathbb{R})$$

such that

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

This gives us

$$\begin{bmatrix} a & a+b & b+c \\ d & d+e & e+f \\ g & g+h & h+i \end{bmatrix} = \begin{bmatrix} a+d & b+e & c+f \\ d+g & e+h & f+i \\ g & h & i \end{bmatrix}$$

Comparing each element, we get  $d = g = h = 0$ ,  $e = a$ ,  $f = b$  and  $i = a$ . Therefore the centralizer is

$$\left\{ \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix} : a \neq 0 \right\}$$

(e) Let

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in GL_3(\mathbb{R})$$

such that

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

This gives us

$$\begin{bmatrix} c & a & b \\ f & d & e \\ i & g & h \end{bmatrix} = \begin{bmatrix} d & e & f \\ g & h & i \\ a & b & c \end{bmatrix}$$

Comparing each element, we get  $d = f = g = 0$  and  $e = g$ . Therefore the centralizer is

$$\left\{ \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix} : a^3 + b^3 + c^3 - 3abc \neq 0 \right\}$$

---

### Result

(a)

$$\left\{ \begin{bmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{bmatrix} : a \neq 0, e \neq 0, i \neq 0 \right\}$$

(b)

$$\left\{ \begin{bmatrix} a & b & 0 \\ d & e & 0 \\ 0 & 0 & i \end{bmatrix} : i \neq 0, ae - bd \neq 0 \right\}$$

(c)

$$\left\{ \begin{bmatrix} a & b & c \\ 0 & a & 0 \\ 0 & h & i \end{bmatrix} : a \neq 0, i \neq 0 \right\}$$

(d)

$$\left\{ \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix} : a \neq 0 \right\}$$

(e)

$$\left\{ \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix} : a^3 + b^3 + c^3 - 3abc \neq 0 \right\}$$

12. a

The class equation is given by

$$|G| = |C_1| + |C_2| + \cdots + |C_k|$$

For  $n = 3$ , we have  $|G| = |C_1| + |C_2| + |C_3|$ . Since the conjugacy class of identity element consist of itself so  $|C_1| = 1$ . Also since  $|C_2|$  and  $|C_3|$  divide  $|G|$ , we get

$$1 = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$$

where  $|G| = a$ ,  $|G|/|C_2| = b$  and  $|G|/|C_3| = c$  are positive integers. This has solutions  $a = b = c = 3$  and  $a = 6, b = 2, c = 3$ . Therefore there are two groups with exactly three conjugacy classes. First case, we get  $a = |G| = 3$  so the group is cyclic group isomorphic to  $C_3$ .

For  $|G| = 6$ , we know that dihedral group  $D_3$  has six elements and three conjugacy classes  $\{1\}, \{x, x^2\}$  and  $\{y, xy, x^2y\}$ .

For  $n = 2$ , we get

$$|G| = |C_1| + |C_2|$$

where  $|C_1| = 1$  which is conjugacy class of trivial element. Since  $|C_2|$  divides  $|G|$ , we may write it up as

$$1 = \frac{1}{a} + \frac{1}{b}$$

where  $|G| = a$  and  $|G|/|C_2| = b$  are positive integers. The only integral solution is  $a = 2, b = 2$ . Thus the order of the group is  $|G| = 2$ . This is isomorphic to  $C_2$ .

For  $n = 1$ , the only group with one conjugacy class is trivial group.

## Result

3 of 3

Trivial group,  $C_2, C_3$  and dihedral group  $D_3$ .

## 13. a

Given that  $N$  is normal subgroup of a group  $G$  with odd order. Also it is given that  $|N| = 5$ . The subgroup  $N$  must be cyclic subgroup so let us denote  $N = \{1, x, x^2, x^3, x^4\}$ . Since  $N$  is normal, we get

$$g^{-1}Ng = N \implies g^{-1}xg = x^s$$

for any  $g \in G$  where  $s \in \{1, 2, 3, 4\}$ . If  $x$  is not in the center of the group then  $x^s \in C_x$  for  $x \neq 1$ . Also, we get

$$(g^{-1}xg)^r = g^{-1}x^r g = (x^s)^r$$

Now if  $x^r \in Z$  then  $g^{-1}x^r g = x^r$ . This gives  $x^{sr} = x^r$  i.e.  $sr \equiv r \pmod{5}$  which implies  $s = 1$ . However since  $x \notin Z$ , this is false. Therefore none of the elements lie in the center and also we have

$$g^{-1}x^p g = x^q$$

where  $p \neq q$  and  $p, q \in \{1, 2, 3, 4\}$ . This shows that elements of  $N$  form their own conjugacy class. Since none of the elements are in center, thus the conjugacy class  $C_x$  must contain all four elements or two elements (It cannot contain three elements as it would leave another single element which would be in center).

However, it is given that order of  $G$  is odd. But since we have conjugacy class of order 2 or 4 which contradicts the fact that order of group must be divisible by 2 or 4. Therefore  $x$  must lie along the center of group. Since  $Z$  is normal group itself, it must contain  $N$ .

## Result

Show that if  $N$  were not contained in center  $Z$  then we get conjugacy class of order two or four.

14. a

(a) The class equation of group  $G$  is given as  $1 + 4 + 5 + 5 + 5$ . Let  $|C_x| = 4$  then centralizer of  $x$ ,  $Z(x)$  has order  $|Z(x)| = |G|/|C_x| = 20/4 = 5$ . This shows there exists one subgroup, centralizer of  $x$  i.e.  $Z(x)$ , which is of order 5. Since the order of group is 5, which is prime, the subgroup must be cyclic and each element is of order 5.

Let this group be  $N = \langle r \rangle$ . Now, for  $g \in G$  and suppose  $g^{-1}rg = a \notin N$ . Consider the cyclic group generated by  $a$ . This group must be of order 2 or 4 since there is only one group of order 5. If this is of order 2 then we get

$$g^{-1}r^2g = 1 \implies r^2g = g \implies r^2 = 1$$

If this group is of order 4 then we get

$$g^{-1}r^4g = 1 \implies r^4g = g \implies r^4 = 1$$

Hence  $g^{-1}rg \in N$  so  $N$  must be normal subgroup.

(b) Since there exists conjugacy class of order 5, say  $C_y$ , the centralizer  $Z(y)$ , which have order  $|Z(y)| = |G|/|C_y| = 4$ , is the subgroup of order 4.

There are three such subgroups since there exists  $5 + 5 + 5$  in the class equation. Let these subgroups be  $Z(y_1)$ ,  $Z(y_2)$  and  $Z(y_3)$ . If one of these subgroups were normal, say  $Z(y_1)$  the quotient group  $G/Z(y_1)$  would have order

$$\frac{|G|}{|Z(y_1)|} = \frac{20}{4} = 5$$

which would be cyclic. This implies that  $G$  is abelian group which contradicts that  $G$  has class equation  $1 + 4 + 5 + 5 + 5$ .

## Result

3 of 3

(a) By class equation, there is one subgroup of order 5 which is cyclic. Let  $r$  be generator then consider the group generated by  $g^{-1}rg$ . This shows the subgroup is normal. (b) By class equation, there exists three conjugacy classes of order 5 so there must be subgroups (centralizer) whose order is 4. This cannot be normal since the order of quotient group by this subgroup would be 5 (which would imply that quotient group is cyclic) which implies  $G$  is abelian.

15. a



The order of  $GL_n(\mathbb{F}_q)$  is given by

$$|GL_n(\mathbb{F}_q)| = \prod_{k=0}^{n-1} (q^n - q^k)$$

In our case,  $n = 2$  and  $q = 3$  which gives us  $|GL_2(\mathbb{F}_3)| = 48$ . The center of this groups is

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \right\}$$

The conjugacy class of

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \right\}$$

The conjugacy class of

$$\begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 0 \end{bmatrix} \right\}$$

The conjugacy class of

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix} \right\}$$

The conjugacy class of

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \right. \\ \left. \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \right\}$$

The conjugacy class of

$$\begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \right\}$$

The conjugacy class of

$$\begin{bmatrix} 0 & 2 \\ 1 & 2 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} 0 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix} \right\}$$

## Step 2

2 of 3

Thus, the class equation of  $GL_2(\mathbb{F}_2)$  is given by  $1 + 1 + 6 + 6 + 6 + 8 + 8 + 12$ . First  $1 + 1$  is for centralizer and it has three conjugacy classes of order 6, two conjugacy classes of order 8 and one of order 12.

## Result

3 of 3

The class equation of  $GL_2(\mathbb{F}_2)$  is given by  $1 + 1 + 6 + 6 + 6 + 8 + 8 + 12$ .

16. a

Let  $G, G'$  be two finite groups, and assume that  $\phi : G \rightarrow G'$  is a surjective group homomorphism. Let  $C$  be a conjugacy class of  $G$ . Let  $x \in C$ , and assume that  $C'$  is the conjugacy class of  $\phi(x)$  in  $G'$ . First we prove that  $\phi$  maps  $C$  surjectively onto  $C'$ . Let  $y \in C'$ , so there exists  $g \in G$  such that  $y = gxg^{-1}$ . Then  $\phi(y) = \phi(gxg^{-1}) = \phi(g)\phi(x)\phi(g)^{-1}$ . So  $\phi(y) \in C'$ . This proves that  $\phi(C) \subset C'$ . Let  $z \in C'$ , and so there exists  $k \in G'$  such that  $z = k\phi(x)k^{-1}$ . Now, since  $\phi$  is surjective, there exists  $h \in G$ , such that  $\phi(h) = k$ . Then,  $\phi(hxh^{-1}) = \phi(h)\phi(x)\phi(h)^{-1} = k\phi(x)k^{-1} = z$ , but  $h x h^{-1} \in C$ . So, we have shown that  $z \in C'$ , then there exist an element  $m = h x h^{-1}$ , such that  $\phi(m) = z$ . We conclude that  $\phi$  maps  $C$  onto  $C'$ .

## Step 2

2 of 3

Now, we show that  $|C'| \mid |C|$ . First observe that by first isomorphism theorem,  $G/\text{Ker}\phi \cong G'$  and hence  $|G'| \mid |G|$ . Let  $C(x), C(\phi(x))$ , denote the respective centralizer in their respective groups. We know,  $|C| = \frac{|G|}{|C(x)|}$ , similarly,  $|C'| = \frac{|G'|}{|C(\phi(x))|}$ . Now, observe that  $\phi(C(x)) \subset C(\phi(x))$ . Indeed, if  $t \in C(x) \implies tx = xt$ . Now,  $\phi(t)\phi(x) = \phi(tx) = \phi(xt) = \phi(t)\phi(x)$ . So,  $\phi(t) \in C(\phi(x))$ . Now, since  $C(x)$ , is a subgroup, so we can restrict  $\phi$  on  $C(x)$ , to get the restriction map  $\psi : C(x) \rightarrow C(\phi(x))$ . Again, due to first isomorphism theorem, we have  $|C(\phi(x))|$  divides  $|C(x)|$ . So  $\frac{|C|}{|C'|} = \frac{|G|}{|G'|} \cdot \frac{|C(\phi(x))|}{|C(x)|}$ . Since  $|G'| \mid |G|$ , and  $|C(\phi(x))|$  divides  $|C(x)|$ , it is clear that either  $|C| \mid |C'|$  or vice versa. But, since from the above part,  $\phi(C) = C'$ , we have  $|C'| \leq |C|$ . Hence, we conclude that  $|C'|$  divides  $|C|$ .

## Result

3 of 3

I have solved the problem in two parts. First part proves the  $C$  maps surjectively to  $C'$  under  $\phi$ . The basic idea is to use the definition of homomorphism of groups and the definition of conjugacy class. The second part proves the divisibility, by using first isomorphism theorem, and other elementary arguments.

17. a

Let  $G$  be a group of order  $pq$ , where  $p, q$  are primes. We have to use class equation to show that  $G$  has an element of order  $p$ . Assume, that  $G$  doesn't have an element of order  $p$ . Now, by Lagrange's theorem, the possible orders of non-identity elements are  $q$  or  $pq$ . Now, if there is an element of order  $pq$  say  $x$ , then observe that  $x^q$  has order  $p$ , which is a contradiction to our assumption. Now, we have that all non-identity element have order  $q$ . Now, consider  $Z(G)$ . Now,  $|Z(G)|$  can be  $1, q$ , or  $pq$ . Observe that  $|Z(G)| = pq$  implies that there exists an element in  $Z(G)$ , and hence in  $G$ , of order  $p$ . So, we have the above possibilities. Now, if  $|Z(G)| = q$ , then  $G/Z(G)$  is cyclic, and hence  $G$  is abelian, contradicting the fact that  $|Z(G)| = q$ . So, we have  $|Z(G)| = 1$  or  $pq$ . We deal with each of these cases separately.

Case 1: Assume  $|Z(G)| = 1$ . Then  $|C(x)| = q$  for each non-identity  $x \in G$ . Observe  $|C(x)| \neq p$ , for each  $x$ , as it contradicts the initial assumption. Let  $C_1, C_2 \dots C_{k+1}$ , be the conjugacy classes, with  $C_1 = \{1\}$ . From class equation, we get  $1 + pk = pq$ , as  $\frac{|G|}{|C(x)|} = p$ , for every non-identity  $x \in G$ . But, now this implies  $p \mid pq - 1$ , which is a contradiction. So, case 1 is contradicted as a whole.

Case 2: Assume  $|Z(G)| = pq$ . This implies  $G$  is abelian. Let  $H = \langle x \rangle$ , where  $x$  is an element of order  $q$ . Then  $|H| = q$ . This  $G/H$  has order  $p$ , and hence has an element of order  $p$ . Let  $tH$  has order  $p$ . So,  $t^p \in H$ . Now,  $t^p \neq 1$ , and hence  $t^p$  has order  $q$ . But then  $t$  has order  $pq$ , which is again a contradiction. So, we conclude case 2 is also not possible.

Finally, since all the possibilities have been exhausted, and we have contradicted all such possibilities, due to the initial condition that we assumed. Hence, we conclude that the initial condition must be wrong, and hence there always exists an element of order  $p$ .

## Result

2 of 2

As has been asked in the question, I have only used the class equation formula to prove the result. The proof is by contradiction, by assuming that there is no element of order  $p$ , and then contradicting each possible cases arising out of it.

18. a

Consider the following matrices:

$$A = \begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix}, B = \begin{bmatrix} 0 & -1 \\ 1 & d \end{bmatrix}$$

Determine which pairs of matrices are conjugate elements of following group.

(a)

Consider the following group:

$$GL_2(\mathbb{R})$$

Consider the permutation matrix:

$$P = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

To show that  $PAP^{-1} = B$ ,

$$\begin{aligned} PAP^{-1} &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & d \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & d \end{bmatrix} \end{aligned}$$

$$= B$$

Further,

$$\begin{aligned} \det(P) &= -1 \cdot 1 - 0 \\ &= -1 \end{aligned}$$

So  $P \in GL_2(\mathbb{R})$ , this shows that  $A$  and  $B$  are conjugate as elements of  $GL_2(\mathbb{R})$ . That is pairs of matrices are conjugate elements of  $GL_2(\mathbb{R})$ .

(b)

Consider the following group:

$$SL_2(\mathbb{R})$$

Now suppose they were conjugate by some element  $Q \in SL_2(\mathbb{R})$  say  $QAQ^{-1} = B$ , where

$$Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then,  $QA = BQ$

First calculate  $QA$  and  $BQ$  as shown below,

$$\begin{aligned} QA &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix} \\ &= \begin{bmatrix} -b & a+bd \\ -d & c+d^2 \end{bmatrix} \end{aligned}$$

And,

$$\begin{aligned} BQ &= \begin{bmatrix} 0 & -1 \\ 1 & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\ &= \begin{bmatrix} -c & -d \\ a+cd & b+d^2 \end{bmatrix} \end{aligned}$$

So,

$$\begin{bmatrix} -b & a+bd \\ -d & c+d^2 \end{bmatrix} = \begin{bmatrix} -c & -d \\ a+cd & b+d^2 \end{bmatrix}$$

Therefore,

$$-b = -c, a+bd = -d, -d = a+cd, c+d^2 = b+d^2$$

This implies that,

$$b = c$$

$$a+bd = a+cd$$

This implies that,

$$b = c$$

And,

$$c+d^2 = b+d^2$$

This implies that,

$$c = b$$

And no restriction on  $a$  and  $d$ .

Suppose  $a$  and  $d$  are zero, then, matrix  $Q$  is,

$$Q = \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix}$$

So,

$$\det(Q) = -c^2 < 0$$

This is contradicting that  $Q \in SL_2(\mathbb{R})$

So,  $A$  and  $B$  are not conjugate as elements of  $SL_2(\mathbb{R})$ . that is **pairs of matrices are not conjugate elements of  $SL_2(\mathbb{R})$ .**

## Section 3

### 1. a

#### Fixed point theorem:

Let  $G$  be a  $p$ -group and let  $S$  be a finite set on which  $G$  operates. If the order of  $S$  is not divisible by  $p$ , there is a fixed point for the operation of  $G$  on  $S$ -an element's' whose stabilizer is the whole group.

**Proof of fixed point theorem:**

Assume that there is no fixed point for the operation of  $G$ .

Since  $G$  is a  $p$ -group one can write,

$$|G| = p^e$$

Also here  $S$  is the stabilizer of the whole group then one can write,

$$|G| = s$$

Thus from the above two equations one can equate,

$$s = p^e$$

This implies that, each element in the finite set  $S$  is divisible by  $p$ .

Thus,

The order of  $S$  is divisible by  $p$ .

This is a contradiction to the given statement that the order of  $S$  is not divisible by  $p$ .

Thus, the assumption that there is no fixed point is wrong.

**Hence**, there is a fixed point for the operation of  $G$  on  $S$ .

## 2. a

Suppose  $G$  is a group, and  $Z$  is its center. Given that:  $G/Z$  is cyclic. To prove:  $G$  is abelian and hence  $G = Z$ .  
 Now, since  $G/Z$  is cyclic,  $G/Z = \langle gZ \rangle$ . Let  $x, y \in G$ . Consider  $xZ, yZ \in G/Z$ . Then,  $xZ = (gZ)^m = g^mZ$ , and,  $yZ = g^mZ = g^nZ$ , for some positive integer  $n, m$ . So, we can write  $x = g^mh, y = g^nk$ , for some  $h, k \in Z$ .  
 Now,  $xy = g^mhg^nk = g^mg^nhk = g^{m+n}hk = g^ng^mhk = g^ng^mkh = g^nk g^mh = yx$ . Note that, in each step of the equality we have used the fact that  $h, k \in Z$ , and hence they commute with each element of  $G$ . So  $x, y \in G$  then  $xy = yx$ . Therefore,  $G$  is abelian, and hence  $G = Z$ .

## Result

2 of 2

We have proved that if  $G/Z(G)$  is cyclic, then  $G$  must be abelian.

## 3. a

Non-abelian also defined as non-commutative group is defined as the group in which there exists at least one pair of elements  $x$  and  $y$  of  $G$  such that;

$$xy \neq yx$$

[Comment](#)

### Step 2 of 4 ^

Consider a nonabelian group  $G$  whose order is  $p^3$ , where  $p$  is prime.

**a.**

To find: the possible orders of the centre  $Z$

The possible orders for the centre  $Z$  is,

According to the first Sylow theorem,

A finite group whose order is divisible by a prime  $p$ , contains an element of order  $p$

Here, since  $Z$  is the centre of the non Abelian group

The order of the group is a smallest positive integer for which is the identity element

**Therefore, the possible orders of  $z$  is**  $\boxed{p^3}$



b.

Let  $x$  be an element of  $G$  that is not in  $Z$

To find: the order of the centralizer  $Z(x)$

Let,  $x \in G$

And  $x \notin Z$

The centralizer  $Z(x)$  of an element  $x$  of  $G$  contains  $x$ , and it contains the centre  $Z$

The center of an element is the set of elements that commutes with every element of ' $G$ '

Hence,  $Z(x) = 3$

[Comment](#)

---

Step 4 of 4 ^

c.

To find: the possible class equations for  $G$ ;

The class equation can be formed using the solutions,

$$a = \pm 1$$

$$c = 0$$

$$a = 0$$

$$c = \pm 1$$

Here the order of the group  $G$  is 3

Hence, the possible class equation can be written as  $3 = 1 + 1 + 1$

4. a

A group is defined as an algebraic structure consisting of a set of elements combined with an operation that combines any two elements to form the third element.

[Comment](#)

---

Step 2 of 2 ^

To classify: the groups of order 8;

The groups of order 8:

Let  $G$  be the group and  $a \in G$

Thus the possible group with order 8 is,

The order of an element in a group is the smallest positive integer for which is the identity element

In other words;

$$a^m = e$$

Where,  $a^m$  Denotes the products of ' $m$ '

And, ' $e$ ' denotes Identity element

Hence here,

$$a = \{1\}$$

## Section 4

1. a

The icosahedron has 5 triangles around each vertex. It has 20 faces, 12 vertices, and 30 edges.

[Comment](#)

---

Step 2 of 3 ^

Stabilizer is defined below:

Let  $G$  be a permutation group on a set  $\alpha$  and  $x$  be an element of  $\alpha$ . Then;

$$G_x = \{g \in G : g(x) = x\}$$

Then  $G_x$  is called the stabilizer of  $x$  and consists of all the permutations of  $G$  that produce group fixed points in  $x$

[Comment](#)

---

Step 3 of 3 ^

The icosahedral group operates on the set of five inscribed cubes in the dodecahedron. The stabilizer of one of the cubes is:

Since the class equation of the icosahedral group is given as.

$$60 = 1 + 20 + 12 + 12 + 15$$

The divisors of sixty

Therefore, the stabilizer of one of the cubes is  $\boxed{1}$

2. a

Subgroup is defined as the group whose members are all members of another group, both with respect to the same operators.

And, normal subgroups are those subgroups which is invariant under conjugation by members of the group of which it is a part.

[Comment](#)

---

Step 2 of 2 ^

To prove: whether  $A_5$  is the only proper normal subgroup of  $S_5$

For the proof first define a function:

$$\phi: I \rightarrow S_5$$

Where  $I$  refers the simple group and the function  $\phi$  defines an isomorphism from  $I$  to the alternating group  $A_5$

Since, this function is trivial as the order of  $I$  and  $A_5$  is 60 and  $\phi$  is injective,

Hence, the image of  $\phi$  which is isomorphic to  $I$  is  $A_5$

**Therefore,  $A_5$  is the only proper normal subgroup of  $S_5$**

3. a

The icosahedron has 5 triangles around each vertex. It has 20 faces, 12 vertices, and 30 edges.

Centralizer of a subset  $S$  of a group  $G$  is the set of elements of  $G$  that commute with each element of  $S$ .

---

[Comment](#)

---

Step 2 of 2 ^

The centralizer of an element of order 2 of the Icosohedral group is:

Let,

$$x \in G$$

The conjugacy class equation for the element 'x' of order 2 is,

$$2 = 1 + 1$$

**Thus, the centralizer of an element of order 2 is the whole group  $G$  since the class consists of the element 'x' alone**

4. a

A tetrahedral group is the set of all self-isometries of  $\mathbb{R}^3$  that send a particular regular tetrahedron to itself.

---

[Comment](#)

---

Step 2 of 4 ^

a.

To determine: The class equation of the tetrahedral group  $T$

Let the class equation of the tetrahedral group  $T$  is the divisors of 12

Thus, the class equation is given by,

$$12 = 1 + 2 + 4 + 3 + 2$$

b.

To prove: That  $T$  has a normal subgroup of order 4 and no subgroup of order 6

First proving that  $T$  has a normal subgroup of order 4

Let  $T$  has a normal subgroup of order 4 since the group itself is a tetrahedral group

Consider,  $A_4$ :

Since all the four elements form 4 group,

$T$  has a normal subgroup of order 4

---

[Comment](#)

---

Step 4 of 4 ^

Now, proving  $T$  has no subgroup of order 6

Consider  $A_4$ :

Assume that  $A_4$  has a subgroup of order 6.

Since there is no element of order 6 in the group  $A_4$  then the group is not a cyclic group

Thus, the group contains 3 elements of order 2 and one identity element

So, these four elements form a subgroup of order 4 ,

It contradicts the fact that  $A_4$  has a subgroup of order 6

Thus,  $A_4$  does not have a subgroup of order 6

**Therefore,  $T$  have a normal subgroup of order 4 and no subgroup of order 6**

5. a

A group is defined as an algebraic structure of a set of elements with an operation that combines any two elements to form a third element.

---

[Comment](#)

---

Step 2 of 3 ^

a.

The tetrahedral group and the octahedral group are subgroup of the continuous group of 3-dimensional rotations.

To determine: The class equation of the octahedral group  $O$ .

The class equation of the octahedral group is given as,

$$24 = 1 + 8 + 6 + 6 + 3$$

b.

To show: That octahedral group contains two proper subgroups and to show that they are normal and show that there are no others.

Consider the group;

$$p = (12)(34)$$

Here the conjugacy class  $C(p)$  contains the 15 pairs of disjoint transpositions each of which generates a conjugate subgroup of  $H$ .

The counting formula shows that the normalizer  $N(H)$  has order eight

**Hence, these are the only two subgroups of octahedral group**

6. a

(a)

To prove that the tetrahedral group  $T$  is isomorphic to the alternating group  $A_4$ :

It is assumed that  $T$  and  $A_4$  be two groups. A map from  $f: T \rightarrow A_4$  is isomorphic if,  $f$  is a bijection then,  $f(xy) = f(x)f(y) \forall x, y \in G$

[Comment](#)

Step 2 of 4 ^

A function is said to be bijection if it is both one-to-one and onto. Now here the tetrahedral group is both one-to-one and onto and also assume,

$$x = x_1; y = x_2$$

Therefore,

$$\begin{aligned} f(xy) &= f(x_1x_2) \\ &= f(x_1)f(x_2) \\ &= f(x)f(y) \end{aligned}$$

Hence, the conditions are satisfied therefore,  $T$  is isomorphic to the alternating group  $A_4$ .

In other words it can be also said that  $T \cong A_4$ .

To prove that the octahedral group  $O$  is isomorphic to the symmetric group  $S_4$ .

Let  $O$  and  $S_4$  be two groups and there is a map from  $f: O \rightarrow S_4$  is isomorphic if  $f$  is a bijection, that is it satisfied the condition

$$f(xy) = f(x)f(y) \forall x, y \in G$$

Function is said to be bijection if it is both one-to-one and onto. The octahedral group is both one-to-one and onto. It is also assumed that,  $x = x_1; y = x_2$

Therefore,

$$\begin{aligned} f(xy) &= f(x_1x_2) \\ &= f(x_1)f(x_2) \\ &= f(x)f(y) \end{aligned}$$

As the both conditions are satisfied, therefore  $O$  is isomorphic to the symmetric group  $S_4$ .

In other words, it can be also said that  $O \cong S_4$ .

(b)

To find the relation when two tetrahedral can be inscribed into a cube  $C$ .

If from the group of rotational symmetries of a tetrahedron is  $A_4$  and the full group of symmetries of a tetrahedron is  $S_4$ . It is concluded that  $A_4$  is a subset of  $S_4$ .

Hence, they are relating with each other as they are **subset** of each other.

7. a

To show that  $G$  has a proper normal subgroup.

[Comment](#)

Step 2 of 3 ^

Let ' $G$ ' be the group of order ' $n$ '.

Let ' $H$ ' be the group of order ' $r$ '.

The condition is,  $n > r!$

That is,

$$\begin{aligned} |G| &= n \\ &> |H| \\ &= r! \end{aligned}$$

Now, to show that ' $G$ ' has a proper normal sub group, it is enough to prove that there exists a normalizer.

That is, for every  $g \in G$ ,

$$N(H) = \{gHg^{-1} = H\}$$

Here, since ' $n$ ' is an order for the group ' $G$ ' there exist the cyclic group and hence there exist an inverse.

Hence,

$$gHg^{-1} \in G \quad \forall H$$

Thus it is proved that,  $G$  has a proper normal subgroup.

**Hence proved**

8. a

(a)

Consider the provided statement to explain about the center of the group. As it is given that the centralizer  $Z(x)$  of the group element ' $x$ ' has order 4.

As it is known that, the stabilizer of an element  $x$  of  $G$ , then it is denoted by  $Z(x)$ .

$$Z(x) = \{x \in G \mid xa = ax, \exists x\}$$

If  $|G| \leq 5$  then  $G$  is abelian. If they follow commutative property then both center and centralizer is same. An element  $x$  of  $G$  is in the center if and only if its centralizer  $Z(x)$  is the whole group  $G$ .

[Comment](#)

Step 2 of 2 ^

(b)

Consider the provided statement to explain about the center of the group. As it is given that the conjugacy  $C(y)$  of the group element ' $y$ ' has order 4.

Let  $y \in G$ . Here,  $C(y)$  is the whole group

Note that, an element ' $x$ ' of  $G$  is in the center if and only if the conjugacy class  $C(y)$  consists of the element  $y$  alone. Hence, the center element must be inside the conjugacy class.

9. a



Let 'x' be an element of the group  $G$  and also it is given that, the centralizer  $Z(x)$  has order  $pq$  where  $p$  and  $q$  are primes.

The objective is to prove that  $Z(x)$  is abelian.

Let,  $x \in G$  and  $x \neq e$ , where  $e$  is the identity element.

Order of  $Z(x)$  is  $pq$ , this implies that  $|Z(x)| = pq$

By the definition of centralizer, we have  $Z(x) = \{y \in G \mid xy = yx\}$

Show that  $Z(x)$  is abelian group.

**Case (i)**

If  $p = q$  then  $|Z(x)| = p^2$

Notice that any group of order  $p^2$ , where  $p$  is prime is abelian group.

Therefore,  $Z(x)$  is abelian.

**Case (ii)**

For  $p \neq q$

Since  $x \in Z(x)$  and  $x \neq e$ , the order of  $x$  must be greater than 1.

That is,  $|x| > 1$

We know that order of an element divides the order of the group.

So, the possible orders for  $x$  is  $p, q$  or  $pq$

That is,  $|x| = p, q$  or  $pq$

**Subcase (i)**

If  $|x| = pq$  then order of an element is equals to the order of group  $Z(x)$ , which means  $Z(x)$  is cyclic.

Every cyclic group is abelian, so  $Z(x)$  is abelian.

**Subcase (ii)**

Assume that  $|x| = p$

By Cauchy's theorem there is an element of order  $q$  say  $y$ .

That is,  $|y| = q$  and  $y \in Z(x)$

It is enough to show that  $|xy| = pq$

Assume  $p < q$ , order of  $xy$  can be one of the among  $p, q$  or  $pq$

Consider  $(xy)^p$

$$\begin{aligned} (xy)^p &= x^p y^p \\ &= ey^p \quad (\text{Since } |x| = p) \\ &= y^p \\ &\neq e \end{aligned}$$

Therefore, the order of  $xy$  is not equals to  $p$

Now consider  $(xy)^q$

$$\begin{aligned} (xy)^q &= x^q y^q \\ &= x^q e \quad (\text{Since } |y| = q) \\ &= x^q \\ &\neq e \end{aligned}$$

Therefore, the order of  $xy$  is not equals to  $q$

Hence, we have only possibility of order  $xy$  is  $pq$

That is,  $|xy| = pq$

This implies that as  $xy \in Z(x)$

So  $Z(x) = \langle xy \rangle$  is cyclic group generated by  $xy$

Hence,  $Z(x)$  is an abelian group.

## Section 5

1. a

(a)

Consider the transpositions:

$$(12)(23)\dots(n-1, n)$$

To show the symmetry for this transpositions relabel the indices as follows:

Let,

$$\phi: I \rightarrow L$$

Denotes the relabeling map that goes from the set of indices to the set  $L$  of letters,

$$\phi(1) = a; \phi(2) = b; \phi(3) = c; \phi(4) = d \dots \text{etc}$$

Hence the relabeled permutation is,

$$\phi \circ p \circ \phi^{-1}$$

Now one can use permutation 'q' of the indices to relabel in the same way.

Thus the result the conjugate,

$$p' = qpq^{-1}$$

Will be a new permutation of the same cycle of indices, For example,

If  $q = (1452)$

The relabeled set will be,

$$\begin{aligned} qpq^{-1} &= (1452) \circ (134)(25) \circ (2541) \\ &= (435)(12) \\ &= p' \end{aligned}$$

Thus, whenever the transpositions are relabeled, their indices are not changed and it follows the symmetry.

**Hence proved**

(b)

Consider the cycle:

$$(123\dots n)$$

Here one can need ' $n-1$ ' number of transpositions to rewrite the cycle because

The cycle can be written as,

$$(123\dots n) = (12)(2,3)\dots(n-1,n)$$

Therefore, one can need product of  $\boxed{n-1}$  transposition.

[Comment](#)

Step 3 of 3 ^

(c)

Consider the cycles:

$$(123\dots n) \text{ And } (12)$$

To prove that cycles  $(123\dots n)$  and  $(12)$  generate the symmetric group  $S_n$ :

Now,

Here the transposition for the first cycle follows the symmetry.

Now the second cycle is the subset of the first cycle.

Since the first cycle is symmetry,

Then it generates the other cycle to form a symmetric group  $S_n$ .

**Hence proved**

2. a

Find centralizer of the element  $(1\ 2)$  in  $S_5$ .

[Comment](#)

Step 2 of 2 ^

The centralizer of the element,

$$(1\ 2)$$

The centralizer consists of elements  $a$  such that  $(1\ 2)a = a(1\ 2)$ .

This equation is possible if either all cycles in  $a$  do not intersect  $(1\ 2)$  or one of the cycles coincides with  $(1\ 2)$ .

In the first case: 6 permutations of 3, 4, 5, and in the second case: same elements multiplied by  $(1\ 2)$ .

Therefore, in total, the centralizer has  $\boxed{12}$  elements.

3. a

To determine the orders of the elements of the symmetric group  $S_7$ :

[Comment](#)

Step 2 of 6 ^

The group of  $S_7$  has 7 elements so consider the cycle  $(1234567)$  has length 7.

Consider the all cycles belongs to symmetric group  $S_7$  and order of element as shown below,

Cycle  $\{e\} \in S_7$  and has order of element is 1.

Cycle  $(1,2) \in S_7$  and has order of element is 2.

Cycle  $(1,2,3) \in S_7$  and has order of element is 3.

[Comment](#)

Step 3 of 6 ^

Consider the 4 elements in group  $S_7$ :

Cycle  $(1,2,3,4) \in S_7$  and has order of element is 4.

Cycle  $(1,2)(3,4) \in S_7$  and has order of element is  $\text{lcm}(2,2) = 2$ .

Consider the 5 elements in group  $S_7$ :

Cycle  $(1,2,3,4,5) \in S_7$  and has order of element is 5.

Cycle  $(1,2,3)(4,5) \in S_7$  and has order of element is  $\text{lcm}(3,2) = 6$ .

[Comment](#)

Step 5 of 6 ^

Consider the 6 elements in group  $S_7$ :

Cycle  $(1,2,3,4,5,6) \in S_7$  and has order of element is 6.

Cycle  $(1,2,3,4)(5,6) \in S_7$  and has order of element is  $\text{lcm}(4,2) = 4$ .

Cycle  $(1,2)(3,4)(5,6) \in S_7$  and has order of element is  $\text{lcm}(2,2,2) = 2$ .

Cycle  $(1,2,3)(4,5,6) \in S_7$  and has order of element is  $\text{lcm}(3,3) = 3$ .

Consider the 7 elements in group  $S_7$ :

Cycle  $(1,2,3)(4,5)(6,7) \in S_7$  and has order of element is  $\text{lcm}(3,2,2) = 6$ .

Cycle  $(1,2,3,4,5)(6,7) \in S_7$  and has order of element is  $\text{lcm}(5,2) = 10$ .

Cycle  $(1,2,3,4)(5,6,7) \in S_7$  and has order of element is  $\text{lcm}(4,3) = 12$ .

Cycle  $(1,2,3,4,5,6,7) \in S_7$  and has order of element is 7.

**Therefore,** the orders of the elements of the symmetric group  $S_7$  are:

$\boxed{1, 2, 3, 4, 5, 6, 7, 10, 12}$ .

4. a

The centralizer of a subset  $S$  of a group  $G$  is the set of elements of  $G$  that commutes with each element of  $S$ .

[Comment](#)

Step 2 of 3 ^

Consider the permutation,

$$\sigma = (153)(246)$$

Now the centralizer  $Z$  is the set of element that commute with every  $\sigma$

Further, the number of conjugates in  $S_7$  for permutation  $(153)(246)$  is;

$$\begin{aligned}\sigma &= \frac{7 \times 6 \times 5}{3} \times \frac{4 \times 3 \times 2}{3} \\ &= 560\end{aligned}$$

And, the number of elements in  $S_7$  is;

$$\begin{aligned}7! &= 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 \\ &= 5040\end{aligned}$$

Now, the centralizer is;

$$\begin{aligned}|S_7 : C_{S_7}| &= \frac{5040}{560} \\ &= 9\end{aligned}$$

Hence,  $|C_{S_7}| = 9$

Regarding the orders of  $Z(\sigma)$ ,  $C(\sigma)$  for the above symmetric group  $S_7$  one can split them into two cycles of odd and even permutations as 3-cycles and 4-cycles.

**Hence, the order of  $Z(\sigma)$ ,  $C(\sigma)$  are 3 and 4 respectively**

5. a

Permutations is defined as the each of several possible ways in which a set or number of things can be ordered or arranged.

[Comments \(1\)](#)

Step 2 of 3 ^

Let ' $p$ ' and ' $q$ ' be two permutations

To prove: The products of  $pq$  and  $qp$  have cycles of equal sizes

Consider,

$$p = (1452)$$

The relabeled set will be,

$$\begin{aligned}pq p^{-1} &= (1452) \circ (134)(25) \circ (2541) \\ &= (435)(12) \\ &= q'\end{aligned}$$

Now, consider,

$$q = (2345)$$

The relabeled set will be,

$$\begin{aligned} qpq^{-1} &= (2345) \circ (234)(45) \circ (5432) \\ &= (435)(25) \\ &= p \end{aligned}$$

Hence, the product of  $pq$  and  $qp$  have cycles of equal sizes

6. a

Subgroup is defined as the group whose members are all members of the other group, where both subject to the same operations.

[Comment](#)

Step 2 of 2 ^

To find: all the normal subgroups of  $S_4$  of order 4

Subgroups of  $S_4$  with order 4 has following normal:

The subgroup can be written as,

$$S_4 = \{1, 2, 3, 4\}$$

Now the normal subgroups consists of an empty set, the Cartesian products of the set

Hence,

**There are four normal subgroups of  $S_4$**

$$\{e, (12)(34), (13)(24), (14)(23)\}$$

7. a

To prove that  $A_n$  is the only subgroup of  $S_n$  with index 2;

[Comment](#)

Step 2 of 3 ^

Let  $H$  be the sub group of  $A_n$ .

$$N < S_n$$

$$\text{Also, } [S_n : N] = 2$$

Here  $N$  must be normal.

So, if  $N$  contains any cycle then it must contain them all.

Also,

$$\text{If } g, h \in N$$

Thus,

$$g^{-1}N = hN$$

Since there is only one coset of  $N$  different from  $N$  and hence,

$$N = ghN$$

Thus,

$$gh \in N$$



If  $N$  contains 2-cycle it must contain them all since  $N$  is normal.

Also,  $S_n$  is generated by its 2-cycles,

$$N = S_n$$

Hence,  $N$  does not contain any 2-cycles and for any 2-cycles  $s_1, s_2, s_1 s_2 \in N$ .

Thus,

$$N = S_n$$

Since  $A_n$  is defined as the group of permutations as an even number of 2-cycles.

Thus,

It is proved that  $A_n$  is the only subgroup of  $S_n$  with index 2.

**Hence proved**

8. a

Find the integers  $n$  such that there is a surjective homomorphism from  $S_n$  to  $S_{n-1}$ .

[Comment](#)

Step 2 of 4 ^

Consider the symmetric groups:  $S_n, S_{n-1}$

Where  $s_n$  and  $s_{n-1}$  are the two distinct groups that follows the symmetry.

Now, let  $\phi$  be the function from  $\phi: S_n \rightarrow S_{n-1}$  and  $\phi$  be a surjective homomorphism

Note that,

If the function is said to be surjective if and only if the set  $S_n$  generates the set  $S_{n-1}$

Order of  $S_n$  is  $n!$  and order of  $S_{n-1}$  is  $(n-1)!$

Suppose that  $n > n-1 \geq 3$  and  $n \geq 5$ .

Assume that there are no surjective homomorphisms  $\phi: S_n \rightarrow S_{n-1}$ .

Suppose that  $\phi: S_n \rightarrow S_{n-1}$  is such a homomorphism. The kernel of the restriction  $\phi$  is either  $A_n$  or trivial because when  $n \geq 5$  then  $A_n$  is simple.

In the first case:

$$\begin{aligned} (n-1)! &= |S_{n-1}| \\ &\leq \left| \frac{S_n}{A_n} \right| \\ &= \frac{n!}{2} \\ &= 2 \end{aligned}$$

The second case:

$$\begin{aligned} (n-1)! &\geq |A_n| \\ (n-1)! &\geq \frac{n!}{2} \\ (n-1)! &\geq \frac{n(n-1)!}{2} \\ 2 &\geq n \end{aligned}$$

Both of the cases are absurd because  $n > n-1 \geq 3$ .

**Hence,**

The integer is  $\boxed{n > 3}$  such that there is a surjective homomorphism from  $S_n$  to  $S_{n-1}$ .

## 9. a

Consider  $X$  to be the finite set of at least two elements, then the permutations of  $X$  fall into two classes of equal size, that is the even permutation and the odd permutation.

[Comment](#)

Step 2 of 2 ^

Let 'q' be a 3-cycles in  $S_n$

The number of permutation can be calculated using the formula:  $n-1$

Hence, if there are 3-cycles, then the number of permutations is given as,

$$\begin{aligned} n-1 &= 3-1 \\ &= 2 \end{aligned}$$

Thus, there must be two permutations with one odd and one even satisfying;

$$pqp^{-1} = q$$

**Therefore, there must be one even permutations for 'q'**

## 10. a

Two elements  $x, y$  are defined to be conjugate of a group  $G$  then there exists an element  $z$  in  $G$  such that;

$$gag^{-1} = b; g \in G$$

Consider the conjugacy mapping;

$$\sigma: S_4 \rightarrow S_4$$

Then the formula for the conjugacy states that;

$$\sigma(i_1, \dots, i_k) \sigma^{-1} = (i_{\sigma(1)} \dots i_{\sigma(k)})$$

Where,  $i_1, \dots, i_k \in S_4$

This implies the number of partitions will be same as that of the order of conjugacy

Here, the number of conjugacy is 4.

Thus, the number of partitions will be 4

Now, assume  $w, x, y, z$  are the four distinct elements in  $S_4$

Thus, the conjugacy class will be of the form of;

$$\{iz\}, \{(w, x)\}, \{(w, x, y)\}, \{(w, x)(y, z)\} \text{ and } \{(wxyz)\}$$

Now, the sizes of these classes are 1, 6, 8, 3 and 6 respectively

Also;

$$|S_4| = 24$$

Hence, the class equation becomes;

$$\begin{aligned} 24 &= 1 + 6 + 8 + 3 + 6 \\ &= 1 + 3 + 6 + 6 + 8 \end{aligned}$$

**Therefore, the formula for the class equation of  $S_4$  is  $24 = 1 + 3 + 6 + 6 + 8$**

For finding the centralizers consider;

$$w = 1$$

$$x = 2$$

$$y = 3$$

$$z = 4$$

So, the conjugacy classes become;

$$\{4\}, \{(12)\}, \{(123)\}, \{(14)(23)\}, \{(1234)\}$$

Centralizers are basically the elements of any group which commutes with every element of the subgroup of that group.

Now, consider the table of some subgroups of  $S_4$  as given below;

Order	Subgroups
4	$\{e, (12), (34), (12)(34)\}$
6	$\{e, (123), (132), (12), (13), (23)\}$
8	$\{e, (12), (34), (12)(34), (13)(24), (14)(23), (1324), (1423)\}$

Hence, from the above table the centralizers are as follows;

$$C((12)) = \{e, (12), (34), (12)(34)\}$$

$$C((123)) = \{e, (123), (132), (12), (13), (23)\}$$

$$C(e, (14), (23), (14)(23), (12)(34), (13)(24), (1243), (1342))$$

Now, find the class equation for  $S_5$

Consider the conjugacy mapping;

$$\sigma : S_5 \rightarrow S_5$$

Then the formula for the conjugacy states that;

$$\sigma(i_1, \dots, i_k) \sigma^{-1} = (i_{\sigma(1)} \dots i_{\sigma(k)})$$

Where,  $i_1, \dots, i_k \in S_5$

And, preceding the procedure as done for  $S_4$ , the following table with size of conjugacy is obtained;

Representation	(1)	(12)	(12)(34)	(123)	(12)(345)	(12345)	(1234)
Size	1	10	15	20	20	24	30

Also;

$$|S_5| = 120$$

Hence, the class equation becomes;

$$120 = 1 + 10 + 15 + 20 + 20 + 24 + 30$$

$$= 1 + 10 + 15 + 20 + 20 + 30 + 24$$

Therefore, the formula for the class equation of  $S_5$  is  $120 = 1 + 10 + 15 + 20 + 20 + 30 + 24$

Now, centralizers are basically the elements of any group which commutes with every element of the subgroup of that group.

Now, consider the table of some subgroups of  $S_5$  as given below;

Order	Subgroups
4	$\{e, (12), (34), (12)(34)\}$
6	$\{e, (123), (132), (12), (13), (23)\}$
8	$\{e, (12), (34), (12)(34), (13)(24), (14)(23), (1324), (1423)\}$

Hence, from the above table the centralizers are as follows;

$$C((1)) = \{e, (1)\}$$

$$C((12)) = \{e, (12), (34), (12)(34)\}$$

$$C((123)) = \{e, (123), (132), (12), (13), (23)\}$$

$$C((12)(34)) = \{e, (14), (23), (14)(23), (12)(34), (13)(24), (1243), (1342)\}$$

## 11. a

Two elements  $x, y$  are defined to be conjugate of a group  $G$  then there exists an element  $z$  in  $G$  such that;

$$gag^{-1} = b; g \in G$$

[Comment](#)

Step 2 of 8 ^

**a.**

Let  $C$  be the conjugacy class

Let  $p$  be an even permutation in  $S_n$

Consider the first case;

Let  $C$  is a conjugacy class in  $A_n$ :

Since,  $p \in C$

Then,  $C$  is the conjugacy class and  $p$  is an even permutation it follows that

$$p \in A_n$$

Therefore,

$$C \in A_n$$

Consider the second case;

Again consider,  $C_1, C_2$  be a conjugacy class in  $A_n$ :

Now, let

$$p \in C_1, C_2$$

Since ' $p$ ' is an even permutation,

It follows that,

$$p \in C_1 \cup C_2$$

The definition of any centralizer states that;

Centralizer consists of those elements in a group  $G$  of a subgroup  $H$  where each centralizer corresponds with the elements in the subgroup

By using the result of above definition, it can be seen clearly that, the centralizer is the union of conjugacy class.

**Hence, the second case occurs in terms of the centralizer of  $p$**

b.

First consider the case of  $A_4$ , now;

$$\begin{aligned} |A_4| &= 4! \\ &= 24 \end{aligned}$$

But the conjugacy class in  $A_4$  splits into two, so;

$$\frac{24}{2} = 12$$

Also, the number of cycles of length 3 in  $A_4$  is;

$$12 - 4 = 8 \text{ cycles}$$

Since, the centralizer is contained in  $A_4$  and the order is 3.

This implies that the conjugacy class of length 3 contains;

$$\frac{12}{3} = 4 \text{ elements}$$

The centralizer of each of the elements of order 2 is the form of abelian subgroup;

$$C = \{e, (12)(34), (13)(24), (14)(23)\}$$

That is the conjugacy class of order 2 contains;

$$\frac{12}{4} = 3 \text{ elements}$$

So, the conjugacy class of  $A_4$  having the orders 1, 3, 4,

Therefore, the class equation of  $A_4$  is;

$$\boxed{12 = 1 + 3 + 4 + 4}$$

Now, consider the case of  $A_5$ , so;

$$\begin{aligned} |A_5| &= 5! \\ &= 120 \end{aligned}$$

But the conjugacy class in  $A_5$  splits into two, so;

$$\frac{120}{2} = 60$$

Any cycle of length 5 has a centralizer contained in  $S_5$  having order 5, say this be;

$$\begin{aligned} Z(\sigma) &= e, \sigma, \sigma^2, \sigma^3, \sigma^4 \\ &\subset A_5 \end{aligned}$$

Thus;

$$\begin{aligned} |C_\sigma| &= \frac{60}{5} \\ &= 12 \end{aligned}$$

Also, the conjugacy class of a 3-cycle or the product of disjoint transpositions is the same as that of  $S_5$  and  $A_5$ , the classes in  $A_5$  splits into two, and have orders 1, 12, 12, 15, 20

Therefore, the class equation of  $A_5$  is;

$$\boxed{60 = 1 + 12 + 12 + 15 + 20}$$

c.

Consider  $a \in A_n$

Then, the cyclic subgroup;

$$\langle a \rangle \subset A_n$$

[Comment](#)

Step 8 of 8 ^

Let this cyclic subgroup be generated by  $a$  that acts on the set  $S = \{1, \dots, n\}$  and decomposes  $S$  into distinct orbits defined as;

$$O_s = \{a^i s : i \in \mathbb{Z}\}$$

Where,  $s \in S$

Now, for each orbit representation  $s$ , let  $N_s$  be the order of  $a$  where the cycle is define as;

$$C_s = \{s a s a^2 s \dots a^{N_s-1} s\}$$

That is all the adjoint conjugacy classes in  $A_n$ , of odd order say,  $\{a_1, a_3, a_5, \dots, a_n\}$  will be in the form of union of all odd order conjugacy class;

$$A_n = a_1 \cup a_3 \cup a_5 \cup \dots \cup a_n$$

12. a

Class equation of a group is defined as the group of conjugation where the group acts on itself by conjugation.

[Comment](#)

Step 2 of 3 ^

Class equations of  $S_6$ :

One simple way to calculate the class equations is determined by  $n!$

Thus, the required explanation is given the table below;

Therefore, the class equation of  $S_6$  is,

[Comments \(1\)](#)

Step 3 of 3 ^

Class equation of:

One simple way to calculate the class equations is determined by

Thus, the class equation of:

Therefore, the class equation of is

## Section 6

1. a



Consider the subgroup  $B$  of invertible upper triangular matrices in  $GL_n(\mathbb{R})$  and the subgroup  $L$  of invertible lower triangular matrices.

Assume,

$$A_{n \times n} \in L$$

So,

$$A_{n \times n}' \in B$$

Then,

$$\det A_{n \times n}' = \det A_{n \times n}$$

And,

$$\text{trace}(A_{n \times n}') = \text{trace}(A_{n \times n})$$

Therefore,

For some  $G \in GL_n(\mathbb{R})$ ,

$$A_{n \times n}' = G A_{n \times n} G^{-1}$$

Because here,

$$\begin{aligned} \det A_{n \times n}' &= (\det G)(\det A_{n \times n})(\det G^{-1}) \\ &= \det A_{n \times n} \end{aligned}$$

Also,

$$\begin{aligned} \text{trace}(A_{n \times n}') &= \text{trace}(G A_{n \times n} G^{-1}) \\ &= \text{trace}(A_{n \times n}) \end{aligned}$$

Thus,

Every invertible lower triangular matrix is conjugate to its transpose, that is, upper triangular matrix.

As a result,

For some  $G \in GL_n(\mathbb{R})$ ,

$$B = GLG^{-1}$$

Hence, the subgroup  $B$  of invertible upper triangular matrices in  $GL_n(\mathbb{R})$  is conjugate to the subgroup  $L$  of invertible lower triangular matrices.

## 2. a

Consider  $B$  be the subgroup of  $GL_n(\mathbb{C})$  of invertible upper triangular matrices.

And,

Let  $U \subset B$  be the set of upper triangular matrices with diagonal entries 1.

Since,

$$N(U) = \{g \in G : gU = Ug\}$$

Let,

$$b \in B$$

Then,

$$bU = Ub$$

Because,

$$\det(bU) = \det(Ub) = \det U$$

And,

The multiplication of upper triangular matrices is upper triangular matrix.

Therefore,

$$N(U) = B$$

Now,

$$N(B) = \{g \in G : gB = Bg\}$$

As,

$$bB = Bb$$

Because,

$$\det(bB) = \det(Bb) = \det B$$

Therefore,

$$N(B) = B$$

Hence,  $N(U) = B$  and  $N(B) = B$ .

---

### 3. a

Consider  $P$  denote the subgroup of  $GL_n(\mathbb{R})$  consisting of the permutation matrices.

Assume,

$$A \in P(\mathbb{R})$$

Let,

$$B = \begin{bmatrix} x & y & \cdots & y \\ y & x & \cdots & \vdots \\ \vdots & \vdots & \ddots & y \\ y & \cdots & y & x \end{bmatrix}_{x \neq y}$$

As,

The matrix  $B$  fixes all the diagonal elements to one element and other all elements besides diagonal to other element.

And,

Both the elements are not equal.

Since,

$$A \begin{bmatrix} x & y & \cdots & y \\ y & x & \cdots & \vdots \\ \vdots & \vdots & \ddots & y \\ y & \cdots & y & x \end{bmatrix}_{x \neq y} A^{-1} = \begin{bmatrix} a & b & \cdots & b \\ b & a & \cdots & \vdots \\ \vdots & \vdots & \ddots & b \\ b & \cdots & b & a \end{bmatrix}_{a \neq b}$$

For some,

$$\begin{bmatrix} a & b & \cdots & b \\ b & a & \cdots & \vdots \\ \vdots & \vdots & \ddots & b \\ b & \cdots & b & a \end{bmatrix}_{a \neq b} \in B$$

Therefore, by definition of normalizer,

$$N(P) = \left\{ x, y \in \mathbb{R} : \begin{bmatrix} x & y & \cdots & y \\ y & x & \cdots & \vdots \\ \vdots & \vdots & \ddots & y \\ y & \cdots & y & x \end{bmatrix}_{x \neq y} \right\}.$$

4. a

Consider  $H$  be a normal subgroup of prime order  $p$  in a finite group  $G$ .

And  $p$  is the smallest prime that divides the order of  $G$ .

[Comment](#)

Step 2 of 2 ^

Since,

The subgroup  $H$  be a normal subgroup of prime order  $p$  in a finite group  $G$ ,

Then,

The order of  $H$  will divide the order of  $G$ .

But,

Center of  $G$  is also a normal subgroup of  $G$ .

So,

The order of  $Z(G)$  will also divide the order of  $G$ .

As,

The prime number  $p$  is the smallest prime that divides the order of  $G$ .

Therefore,  **$H$  is in the center of  $G$ .**

5. a

Consider  $p$  be a prime integer and let  $G$  be a  $p$ -group.

And,

Let  $H$  be a proper subgroup of  $G$ .

Since  $G$  be a  $p$ -group.

Let order of  $G$  be  $p^n$ .

So, the order of group  $G$  is  $p^k, k < n$ .

And,

Let, the subgroup  $H$  is not center of group  $G$ .

Then,

The normalizer  $N(H)$  will contain center of group  $G$ .

As,

The normalizer of center of a group  $G$  is equal to center itself. Therefore, the normalizer  $N(H)$  of  $H$  is strictly larger than  $H$ .

Because,

The normal subgroup of index  $p$  is largest proper normal subgroup of  $G$ .

So,  $H$  is contained in that normal subgroup of index  $p$ .

Hence, **the normalizer  $N(H)$  of  $H$  is strictly larger than  $H$  and that  $H$  is contained in a normal subgroup of index  $p$ .**

6. a

Let  $G$  be a group. Let  $H$  be a proper subgroup of  $G$ . The normalizer of  $H$  in  $G$  is the set of all elements of  $G$  which commute with  $H$ .

Mathematically,

$$N(H) = \{x \in G \mid xH = Hx\}$$

Orbit stabilizer theorem-For a group  $G$  which acts on a finite set  $X$  following relation holds

$$|Orb(x)| = [G : Stab(x)] = \frac{|G|}{|Stab(x)|}$$

(a)

Let  $G$  be a finite group with order  $n$ ,

Let  $H$  be a proper subgroup, let  $[G : H] = m > 1$ .

Let  $N(H)$  be the normalizer of  $H$  in  $G$ , which contains  $H$ .

Then  $[G : N(H)] \leq [G : H]$

Let  $G$  act by conjugation on the set  $H$

Mathematically let  $g \in G$  and  $x \in H$  be arbitrary elements define the group action by

$$(g, x) \rightarrow gxg^{-1}$$

Now evaluate orbit of  $H$

Let  $x \in H$  be arbitrary

Then,  $Orb(x) = \{y = gxg^{-1} \mid g \in G\}$

$$Orb(x) = \{yg = gx \mid g \in G\}$$

So the orbit of  $H$  is the set of all conjugate subgroups.

The stabilizer of  $H$  is given by the set

$$Stab(x) = \{g \in G \mid x = gxg^{-1}\}$$

Thus stabilizer of  $H$  is the set  $N(H)$ .

So by the Orbit-Stabilizer Theorem,

The number of all conjugate subgroups is equal to  $[G : N(H)]$

Since each of the conjugate subgroups has cardinality equal to that of  $H$ , and each contains the identity element  $e$

So there are at most  $1 + [G : N(H)](|H| - 1)$  elements in the union.

Since  $m > 1$

Hence,

$$\begin{aligned} 1 + [G : N(H)](|H| - 1) &\leq 1 + [G : H](|H| - 1) \\ 1 + [G : H](|H| - 1) &= 1 + |G| - m \\ &= |G| + (1 - m) < |G| \end{aligned}$$

So the union of the conjugate subgroups is a proper subset.

**Therefore, the union of all conjugate subgroups of  $H$ , where  $H$  is a proper subgroup of finite group  $G$  is not equal to the group  $G$ .**

(b)

Let  $G$  be a finite group with order  $n$ . Let  $H$  be a proper subgroup of  $G$ .

Assume on contrary that  $H$  intersects every conjugacy class of  $G$ .

Then the following statement holds

$$G = \bigcup_{g \in G} gHg^{-1}.$$

Let  $g_1$  and  $g_2$  lie in the same co-set of  $G/H$

$$\text{Thus } g_1Hg_1^{-1} = g_2Hg_2^{-1}$$

$$\text{So } G = \bigcup_{g \in G} gHg^{-1} \text{ can be written as } G = \bigcup_{g \in G/H} gHg^{-1}$$

Hence the number of sets in the above defined union is given by

$$|G/H| = \frac{|G|}{|H|}$$

Also every element of these sets contains exactly  $|H|$  elements.

$$\text{Since } G = \bigcup_{g \in G/H} gHg^{-1}, \text{ so}$$

$$\begin{aligned} |G| &= \left| \bigcup_{g \in G/H} gHg^{-1} \right| \\ &= \left( \frac{|G|}{|H|} \right) |H| \end{aligned}$$

Thus all these must be disjoint otherwise  $G = \bigcup_{g \in G/H} gHg^{-1}$  fails to hold.

But all these sets being subgroups contain identity element.

A contradiction has been obtained. This contradiction arises due to the wrong assumption that  $H$  intersects every conjugacy class of  $G$ .

**Therefore, there exists a conjugacy class  $C$  of the finite group  $G$  which is disjoint from  $H$ .**

## Section 7

1. a

Consider  $n = p^e m$  and let  $N$  be the number of subsets of order  $p^e$  in a set of order  $n$ .

[Comment](#)

Step 2 of 2 ^

Since,

$$n = p^e m$$

And,

Number of subsets of order  $p^e$  in a set of order  $n = N$

The number  $N$  in binomial coefficient is  $\binom{n}{p^e}$ .

And,

The number  $N$  is not divisible by  $p$ .

Also,

$$m - k = p^e - k$$

$$n - k = p^e m - k$$

Therefore, the congruence class of  $N$  modulo  $p$  is  $m$ .

2. a

Consider  $G_1 \subset G_2$  be groups whose orders are divisible by  $p$  and let  $H_1$  be a Sylow  $p$ -subgroup of  $G_1$ .

Since,

$$G_1 \subset G_2$$

And, both orders are divisible by  $p$ .

Also,

The subgroup  $H_1$  is a Sylow  $p$ -subgroup of  $G_1$ .

So,

By conclusion of Sylow theorems,

There exists a Sylow  $p$ -subgroup  $H_2$  of  $G_2$  such that  $H_1 \subseteq H_2$ ,

Because,

$$G_1 \subset G_2$$

As,

$$H_1 \subseteq G_1, G_1 \subset G_2, H_2 \subseteq G_2$$

Thus,

$$H_1 \subseteq H_2 \cap G_1$$

But,

$$H_2 \cap G_1 \subseteq H_1$$

Therefore,  $H_2 \cap G_1 = H_1$

Hence, there is a Sylow  $p$ -subgroup  $H_2$  of  $G_2$  such that  $H_1 = H_2 \cap G_1$ .

3. a



Consider the elements of order 5 in a group of order 20.

[Comment](#)

Step 2 of 2 ^

By Third Sylow theorem,

The number of its sylow- 5 subgroups divides 4 and is congruent to 1 modulo 5.

Since,

The only such integer is 1.

Therefore,

There is one sylow- 5 subgroup, say  $H$  and it is a normal subgroup.

And,

The subgroup  $H$  is cyclic of order 5.

As,

The order of an element divides the order of the group and identity element of order 1.

Therefore, **there are 4 elements of order 5 in a group of order 20.**

4. a

Consider the groups of order  $pq$  and  $p^2q$  where  $p$  and  $q$  are primes.

(a)

Assume,

The order of a group  $K$  is  $pq$ .

Then,

By Sylow third theorem,

If the number of Sylow  $p$ -subgroups is  $n$  then  $n$  should be a divisor of  $q$  and is congruent to 1 modulo  $p$ .

So,

The possibilities of Sylow  $p$ -subgroups are 1 and  $q$ .

And,

If  $p$  is greater than  $q$  then there will be one Sylow  $p$ -subgroup.

Similarly,

The possibilities of Sylow  $q$ -subgroups are 1 and  $p$ .

Also,

If  $q$  is greater than  $p$  then there will be one Sylow  $q$ -subgroup.

Since  $p$  and  $q$  are primes therefore there will be at least one proper normal subgroup.

Hence, **no group of order  $pq$ , where  $p$  and  $q$  are primes, is simple.**

(b)

Assume,

The order of a group  $K$  is  $p^2q$ .

Then,

By Sylow third theorem,

If the number of Sylow  $p$ -subgroups is  $n$  then  $n$  should be a divisor of  $q$  and is congruent to 1 modulo  $p$ .

So,

The possibilities of Sylow  $p$ -subgroups are 1 and  $q$ .

And,

If  $p$  is greater than  $q$  then there will be one Sylow  $p$ -subgroup.

Now,

The possibilities of Sylow  $q$ -subgroups are 1,  $p$  and  $p^2$ .

Also,

If  $q$  is greater than  $p$  then the number of Sylow  $q$ -subgroup will not be  $p$ ,

If there will be one Sylow  $q$ -subgroup then the group  $K$  will not be simple.

If the number of Sylow  $q$ -subgroups are  $p^2$ ,

Then,

Since each Sylow  $q$ -subgroup has prime order therefore it will be cyclic.

Therefore any two Sylow  $q$ -subgroups will have common element as identity element.

So,

There will be  $(q-1)p^2$  elements of order  $q$ .

Then,

There will be  $p^2$  elements of order  $p$ .

Thus there will be one Sylow  $p$ -subgroup containing  $p^2$  elements.

As there will be at least one proper normal subgroup of group  $K$ ,

Hence, **no group of order  $p^2q$ , where  $p$  and  $q$  are primes, is simple.**

5. a

!!!

6. a

Consider a subgroup of the symmetric group  $S_7$  that is a non-abelian group of order 21.

[Comment](#)

Step 2 of 2 ^

Since,

The order of 7-cycle  $(1\ 2\ 3\ 4\ 5\ 6\ 7)$  is 7.

And,

The order of  $(1\ 4\ 2)(3\ 5\ 6)$  is 3.

Therefore, **the subgroup generated by  $(1\ 2\ 3\ 4\ 5\ 6\ 7)$  and  $(1\ 4\ 2)(3\ 5\ 6)$  will be a non-abelian group of order 21.**

7. a

Let  $G$  be a group with  $|G| = n$  such that  $n = pm$ , where  $p$  is a prime that divides order of the group exactly once

Let  $H$  denote the sylow-  $p$  subgroup of  $G$ .

Let  $S$  denote the set of all sylow-  $p$  subgroup of  $G$

Let  $Q \in S$

Then,  $Q$  be a sylow-  $p$  subgroup of  $G$

Let  $H$  acts on the set  $S$  by conjugation

Define  $O_Q = \text{orbit of } Q \text{ under the conjugation action}$

Then,  $|O_Q|$  divides  $|H|$

Let  $|O_Q| = k$

Then,  $k|p$  since  $|H| = p$

Since  $p$  is a prime number so this implies that

$$k = 1 \text{ Or } k = p$$

Consider the case when  $k = 1$

This implies that  $H$  lies in the normalizer of the subgroup  $Q$

Mathematically,

$$H \subset N(Q)$$

Then,  $H, Q$  are subsets of the normalizer of  $Q$ .

Now from second Sylow theorem,  $H$  and  $Q$  are conjugate to each other

So,  $H = Q$

Now, if  $H \neq Q$  then  $|O_Q| = k = p$

Therefore,  $S$  decomposes into orbits as follows,

The group  $H$  makes its own orbit and rest of the orbit have order  $p$ .

8. a

Let  $GL_n(\mathbb{F}_p)$  denote the group of all invertible matrices of order  $n \times n$  with entries from a finite field  $\mathbb{F}_p$  that contains  $p$  elements.

Let  $\mathbb{F}_p$  be a field such that  $|\mathbb{F}_p| = p$ .

Let  $A = [a_{ij}]_{n \times n}$  be a matrix such that  $\det(A) \neq 0$  and  $a_{ij} \in \mathbb{F}_p$ .

Now the top row of this matrix  $A$  must have at least one non-zero element so that determinant of the matrix is non-zero that is

$$\{a_{1j}\} \neq 0 \text{ for some } j = 1, 2, 3, \dots, n$$

Use combinatorial mathematics to evaluate the total number of such possible ways

So remove the sequence  $\{0, 0, \dots, 0\}$  from the total ways.

So the total number of ways in which this can be done is  $p^n - 1$ .

Now for second row, the second row must not be a multiple of the first row so that the determinant remains non-zero

Hence the total number of possible ways for this to happen is given by

$$p^n - p,$$

where subtraction part denotes the fact that all the  $p$  sequences which are multiple of first row

Now continue this way and evaluate the possible number of ways for the  $j^{\text{th}}$  row which is given by total number of ways minus the  $p^{(j-1)}$  sequences which are the linear combination of all previous rows of the matrix.

Hence the possible number of ways is  $p^n - p^{(j-1)}$ .

Thus the possible number of such matrices  $A$  which satisfy the condition  $\det(A) \neq 0$  is the product of all combination evaluated in above steps.

Mathematically it is given by

$$\prod_{j=1}^n (p^n - p^{(j-1)})$$

So the order of the group  $GL_n(\mathbb{F}_p)$  is  $\prod_{j=1}^n (p^n - p^{(j-1)})$ .

$$\text{Since } |GL_n(\mathbb{F}_p)| = \prod_{j=1}^n (p^n - p^{(j-1)})$$

The above expression can also be written as

$$\prod_{j=1}^n (p^n - p^{(j-1)}) = \prod_{j=0}^{n-1} (p^n - p^j)$$

Take all the possible powers of the number  $p$  common

$$\prod_{j=0}^{n-1} (p^n - p^j) = \left( \prod_{j=0}^{n-1} p^j \right) \left( \prod_{j=0}^{n-1} (p^{n-j} - 1) \right)$$

Assume that  $p$  divides  $p^{n-j} - 1$  for some  $j$ .

Thus there exist  $k \in \mathbb{Z}$  such that

$$\frac{p^{n-j} - 1}{p} = k$$

This implies that  $p^{n-j} = kp + 1$  for some integer  $k$ .

Now under modulo  $p$ ,  $p^{n-j} = kp + 1$  changes as follows

$$p^{n-j} \equiv 1 \pmod{p}$$

But the right hand side of expression is completely divisible by  $p$ .

Hence a contradiction is obtained.

This contradiction arises due to the wrong assumption that  $p$  divides  $p^{n-j} - 1$  for some  $j$ .

Hence,  $p$  does not divide  $p^{n-j} - 1$  for any  $j$

Now the expression  $\prod_{j=0}^{n-1} p^j$  can be written as

$$\prod_{j=0}^{n-1} p^j = p^{\sum_{j=0}^{n-1} j}$$

Since  $\{1, 2, 3, \dots, n-1\}$  is an arithmetic progression so the sum of this arithmetic progression is given by

$$\sum_{j=0}^{n-1} j = \frac{j(j-1)}{2}$$

Put this value in the above derived expression

$$\prod_{j=0}^{n-1} p^j = p^{j(j-1)/2}.$$

So a sylow  $p$ -subgroup of  $GL_n(\mathbb{F}_p)$  has  $p^{j(j-1)/2}$  elements.

Now consider the set  $S(\mathbb{F}_p)$  of all strictly upper triangular matrices that is the set of all the matrices that have diagonal entries as 1, all lower entries are 0 and the upper entries can be any element from the field  $\mathbb{F}_p$ .

Thus the order of the set  $S(\mathbb{F}_p)$  denoted by  $|S(\mathbb{F}_p)|$  is  $p^a$ , where  $a$  denotes the number of entries above the main diagonal.

Now evaluate  $a$

Here  $a$  is given by

$$\begin{aligned} a &= \sum_{j=1}^n n-j \\ &= \sum_{j=0}^{n-1} j \\ &= n(n-1)/2 \end{aligned}$$

So  $|S(\mathbb{F}_p)| = p^{n(n-1)/2}$ , this is equal to the order of the sylow  $p$ -subgroup of  $GL_n(\mathbb{F}_p)$

Thus the set  $S(\mathbb{F}_p)$  is a sylow  $p$ -subgroup of  $GL_n(\mathbb{F}_p)$ .

For the number of sylow  $p$ -subgroup of  $GL_n(\mathbb{F}_p)$ ,

Let  $X$  be the set of all Sylow  $p$ -subgroup of  $GL_n(\mathbb{F}_p)$ .

Let  $GL_n(\mathbb{F}_p)$  act on  $X$  by conjugation. Let  $N$  be the normalizer of the set  $S(\mathbb{F}_p)$ .

Then by counting formula,

$$|GL_n(\mathbb{F}_p)| = |N||X|$$

Now the normalizer of  $S(\mathbb{F}_p)$  is the set of all upper triangular matrices with entries from the field  $\mathbb{F}_p$ .

$$\text{So, } |N| = (p-1)^n p^{n(n-1)/2}$$

Hence the number of elements in  $X$  is given by

$$\begin{aligned} |X| &= \frac{|GL_n(\mathbb{F}_p)|}{|N|} \\ &= \frac{\prod_{j=1}^n (p^n - p^{(j-1)})}{(p-1)^n p^{n(n-1)/2}} \\ &= \prod_{j=1}^n \frac{(p^j - 1)}{p-1} \end{aligned}$$

Therefore, the order of the group  $GL_n(\mathbb{F}_p)$  is  $\prod_{j=1}^n (p^n - p^{(j-1)})$  and the sylow  $p$ -

subgroup of  $GL_n(\mathbb{F}_p)$  is given by the set of all strictly upper triangular matrices with entries from the field  $\mathbb{F}_p$ , also the number of sylow  $p$ -subgroup of  $GL_n(\mathbb{F}_p)$  is

$$\prod_{j=1}^n \frac{(p^j - 1)}{p-1}.$$

9. a

#### SYLOW THEOREM

For every prime factor  $p$  with multiplicity  $n$  of the order of a finite group  $G$ , there exists a Sylow  $p$ -subgroup of  $G$ , of order  $p^n$ .

Given a finite group  $G$  and a prime number  $p$ , all Sylow  $p$ -subgroups of  $G$  are conjugate to each other, i.e. if  $H$  and  $K$  are Sylow  $p$ -subgroups of  $G$ , then there exists an element  $g$  in  $G$  with  $g^{-1}Hg = K$ .

(a)

Let  $G$  be a group such that  $|G| = 33 = 3 \times 11$ .

Let  $s_p$  denote the number of sylow-  $p$  -subgroups of  $G$ .

Evaluate  $s_3$

By Sylow theorem,  $s_3 \equiv 1 \pmod{3}$  and  $s_3$  divides 11

Hence the only possible value for  $s_3$  is 1.

Now evaluate the value of  $s_{11}$

By Sylow theorem,  $s_{11} \equiv 1 \pmod{11}$  and  $s_{11}$  divides 3

Hence the only possible value for  $s_{11}$  is 1.

Use the result that if a group  $G$  has only one sylow-  $p$  -subgroup, then that subgroup is a normal subgroup.

So there exist two normal subgroups of  $G$  say  $P, Q$  with  $|P| = 3$  and  $|Q| = 11$ .

Also,  $P \cap Q = \{e\}$

Hence  $PQ$  is a subgroup of  $G$  that is  $PQ \leq G$ .

But,

$$\begin{aligned} |PQ| &= |P||Q| \\ &= 33 \end{aligned}$$

This implies that  $PQ = G$

Thus,  $G \cong P \times Q$

Since  $P, Q$  are cyclic subgroups of order 3, 11 respectively so,

$$P \cong C_3$$

$$Q \cong C_{11}$$

Thus,  $G \cong C_3 \times C_{11}$

**Therefore, there is only one group upto isomorphism of order 33 which is  $C_3 \times C_{11}$ .**

(b)

Let  $G$  be a group such that  $|G| = 18$ .

Then factorize the order of  $G$ ,

$$\begin{aligned} |G| &= 18 \\ &= 2 \times 3^2 \end{aligned}$$

Let  $s_p$  denote the number of sylow-  $p$  -subgroups of  $G$ .

By Sylow theorem,  $s_3 \equiv 1 \pmod{3}$  and  $s_3$  divides 2

Hence the only possible values for  $s_3$  is 1, that is  $s_3 = 1$ .

So there is only one subgroup of order 9.

Let that subgroup be denoted by  $P$ .

Since  $|P| = 9$ ,

So by Fundamental Theorem of finite Abelian group, either of the following holds

$$P \cong C_9 \text{ Or } P \cong C_3 \times C_3$$

Now for  $s_2$

Since  $s_2$  divides 9 so the  $s_2 \in \{1, 3, 9\}$

Case1- Let  $s_2 = 1$ .

Now if  $Q$  is the only subgroup such that  $|Q| = 2$  then  $Q$  is a normal subgroup of  $G$



Also,

$$\begin{aligned}|PQ| &= |P||Q| \\ &= 18 \\ &= |G|\end{aligned}$$

So,  $G \cong P \times Q$

Thus, if  $P \cong C_3 \times C_3$  then  $G \cong C_3 \times C_3 \times C_2$  and  $G \cong C_3 \times C_6$

If  $P \cong C_9$  then  $G \cong C_9 \times C_2$  and  $G \cong C_{18}$

Case2-Let  $s_2 = 9$

This implies that there are 9 elements of order 2.

Since  $P$  is a subgroup which has 9 elements, so every element which is not in  $P$  is of order 2.

Let  $y \in G - P$ .

Thus  $y$  has order 2.

Since  $[G : P] = 2$ , Then  $G = P \cup Py$

If  $h \in P$  then  $hy$  has order 2.

Mathematically,

$$\begin{aligned}(hy)^2 &= 1 \\ hyhy &= 1 \\ yhy &= h^{-1} \\ yhy &= h^{-1} \\ yh &= h^{-1}y\end{aligned}$$

If  $P \cong C_9$  and  $P = \langle x \rangle$ , then

$$\begin{aligned}G &= P \cup Py \\ &= \{1, x, x^2, \dots, x^8, y, xy, \dots, x^8y\}\end{aligned}$$

The order of the element  $x$  is 9, the order of  $y$  is 2 and  $yx^j = x^{-j}y$ .

Thus  $G \cong D_{18}$ , where  $D_{18}$  denotes the dihedral group of order 18

Now if  $P \cong C_3 \times C_3$ ,

$$P = \{1, x_1, x_1^2, x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_1^2x_2^2\}, \text{ where order of } x_1, x_2 \text{ is } 3 \text{ and they commute.}$$

Hence,

$$\begin{aligned}G &= P \cup Py \\ &= \{1, x_1, x_1^2, x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_1^2x_2^2, y, x_1y, x_1^2y, x_2y, x_2^2y, x_1x_2y, x_1^2x_2y, x_1x_2^2y, x_2x_2^2y\}\end{aligned}$$

Also the following relation holds

$$\begin{aligned}(x_1^i x_2^j y)^2 &= x_1^i x_2^j y x_1^i x_2^j y \\ &= x_1^i x_2^j (x_1^i x_2^j)^{-1} y^2 \\ &= x_1^i x_2^j x_1^{-i} x_2^{-j} \\ &= 1\end{aligned}$$

So every element which is not in  $P$  has order 2.

This group is called the semi-direct product of  $C_3 \times C_3$  with  $C_2$ .

Case3-Let  $s_2 = 3$

Thus there are only 3 elements of order 2.

The elements of  $P$  can have order 1,3,9.

Since order of elements divides the order of group

So all the elements left that is whose order is not 2 and are not in  $P$  must have order 6.

Order of elements cannot be equal to 1,3,9,18.

Since  $1 \in P$ , elements with order 3 would lie in  $P$  and same goes for order 9, finally if there is an element of order 18 then  $G$  is cyclic and hence all subgroups would be normal and it would imply

$$s_2 = 1,$$

This is a contradiction.

Hence there are 9 elements of order 1,3 or 9 in  $P$ , 3 elements of order 2 and 6 elements of order 6.

Let  $y$  be an element of order 2,

So  $y \in G - P$ , and  $G = P \cup Py$

Now if  $P \cong C_9$

Let  $x$  be a generator.

If  $x^i y$  has order 2 then,

$$(x^i y)^2 = 1$$

$$x^i y x^i y = 1$$

$$y x^i y = x^{-i}$$

$$y x^i = x^{-i} y$$

But then,

$$y x^{ki} = y x^i x^{(k-1)i}$$

$$= x^{-i} y x^{(k-1)i}$$

$$= x^{-i} y x^i x^{(k-2)i}$$

$$= x^{-2i} y x^{(k-2)i}$$

Simplify the above expression over and over again to obtain

$$y x^{ki} = x^{-ki} y$$

Also, following holds

$$(x^{ki} y)^2 = x^{ki} y x^{ki} y$$

$$= x^{ki} x^{-ki} y y$$

$$= 1$$

Hence, if  $x^i y$  has order 2, so does  $x^{2i} y, x^{3i} y, \dots$

Since there are only three elements of order 2, they are of the form  $\{y, x^3 y, x^6 y\}$ .

Since  $xy \notin P$  and its order is not 2. Thus  $xy$  has order 6.

But then,  $(xy)^2$  has order 3. But the only elements of order 3 are  $x^3$  and  $x^6$ ,

Thus,

$$(xy)^3 = (xy)^2 xy$$

$$= x^3 xy$$

$$= x^4 y$$

And

$$(xy)^3 = (xy)^2 xy$$

$$= x^6 xy$$

$$= x^7 y$$

But  $(xy)^3$  must have order 2 as  $xy$  has order 6, so  $(xy)^3 \in \{y, x^3y, x^6y\}$

This is a contradiction, thus  $P \cong C_9$  is not the case.

Let  $P \cong C_3 \times C_3$ .

Then there exists an element  $Py$  apart from  $y$  that has order 2.

Let  $x \in P - \{1\}$  be an element such that order of  $xy = 2$ .

This implies that  $yx = x^{-1}y$ , thus  $x^2y$  has order 2, since  $x \in P - \{1\}$  so its order must be 3

So the set  $S = \{1, x, x^2, y, xy, xy^2\}$  is a subgroup and  $S \cong S_3$ .

Now let  $g \in G$  and  $a \in S$  be arbitrary

If  $a$  has order 2, then order of  $gag^{-1}$  is also 2.

But  $S$  contains three elements of order 2 so  $gag^{-1} \in S$ .

If order of  $a$  is not 2,

Then  $a \in P \cap S$ , but  $G = P \cap Py$ , so  $g = y^i h, i \in \{0, 1\}$  and  $h \in P$ .

In this case, since  $y^j = y^{-i}$

$$gag^{-1} = y^i h a h^{-1} y^j$$

Since  $h, a \in P$  and  $P$  is commutative.

So  $hah^{-1} = a$

Thus above expression can be re-written as

$$\begin{aligned} gag^{-1} &= y^i h a h^{-1} y^j \\ &= y^i a y^{-i} \end{aligned}$$

Since  $a, y \in S$ , thus  $gag^{-1} = y^i a y^{-i} \in S$

Thus  $S$  is a normal subgroup.

Now, let  $R$  be the set of subgroups of order 3 in  $G$ .

Then  $|R| = 4$

Thus  $\langle y \rangle = \{1, y\}$  acts on  $R$  by conjugation

So, the orbits must have order 1 or 2 since it must divide  $|\langle y \rangle| = 2$ .

Also, since  $S \triangleleft G$ , the orbit of  $\{1, x, x^2\}$  has only one element.

So,

$$\begin{aligned} 4 &= |R| \\ &= 1 + \sum_{\text{Orbit}} |O| \end{aligned}$$

So, there must be another orbit of order 1

Let  $K$  be this subgroup, and so  $yKy = K$ .

Let  $g \in G$

Then, as before,  $g = y^i h, i \in \{0, 1\}$  and  $h \in P$

Since  $K < P$  and  $P$  is abelian, we have that  $hKh^{-1} = K$

And, since  $yKy = K$ , thus

$$\begin{aligned} gKg^{-1} &= y^i h K h^{-1} y^j \\ &= y^i K y^j \\ &= K \end{aligned}$$

Hence,  $K \triangleleft G$

Thus  $S, K \triangleleft G$  and  $|S||K| = |G|$ , also  $S \cap K = \{1\}$

So,  $G \cong S \times K$

Thus,  $G \cong S_3 \times C_3$

**Therefore there are five groups of order 18 upto isomorphism and they are as follows**

$C_{18}, C_3 \times C_6, D_{18}, S_3 \times C_3$  and the semi direct product of  $C_3 \times C_3$  with  $C_2$ .

(c)

Let  $G$  be a group such that  $|G| = 20$ .

Then factorize the order of  $G$ ,

$$\begin{aligned}|G| &= 20 \\ &= 2^2 \times 5\end{aligned}$$

Then  $s_5 \equiv 1 \pmod{5}$  and  $s_5$  divides 20.

Thus the only possible value of  $s_5$  is 1.

Let  $N$  denote the Sylow-5-subgroup, then this subgroup is normal so  $N \cong C_5$ .

Let  $H \in \text{Sylow-2-subgroup}$ . Then  $|H| = 4$

Clearly,  $G \cong N \rtimes H$

The only possibilities for  $H$  are  $H \cong \mathbb{Z}_4$  and  $H \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ .

Case1- Let  $H \cong \mathbb{Z}_4$

Then  $G \cong N \rtimes_{\phi} \mathbb{Z}_4$ , where  $\phi: \mathbb{Z}_4 \rightarrow \text{Aut}(\mathbb{Z}_5)$  is a homomorphism.

There are four homomorphisms  $\mathbb{Z}_4 \rightarrow \text{Aut}(\mathbb{Z}_5)$  given by

$$1 \rightarrow 0, 1, 2 \text{ or } 3$$

But the homomorphism  $1 \rightarrow 1$  and  $1 \rightarrow 3$  are related by an automorphism which is given by

$$\beta: \mathbb{Z}_4 \rightarrow \mathbb{Z}_4 \text{ such that } \beta(x) = -x.$$

Thus there are three semi-direct products  $\mathbb{Z}_5 \rtimes \mathbb{Z}_4$ ,  $\mathbb{Z}_5 \rtimes_1 \mathbb{Z}_4$  and  $\mathbb{Z}_5 \rtimes_2 \mathbb{Z}_4$

Since  $\text{Center}(\mathbb{Z}_5 \rtimes_1 \mathbb{Z}_4) = \{e\}$  whereas the center of  $\mathbb{Z}_5 \rtimes_2 \mathbb{Z}_4$  has a non-identity element  $2 \in \mathbb{Z}_4$ .

Thus,  $\mathbb{Z}_5 \rtimes_1 \mathbb{Z}_4$  is not isomorphic to  $\mathbb{Z}_5 \rtimes_2 \mathbb{Z}_4$

Case2-Let  $H \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

Then  $G \cong N \rtimes_{\phi} \mathbb{Z}_2 \times \mathbb{Z}_2$

There exists four homomorphisms  $\mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow \text{Aut}(\mathbb{Z}_5)$  which consist of one trivial and the other three are non-trivial homomorphism.

The non-trivial cases are related to each other by the automorphism of  $\mathbb{Z}_2 \times \mathbb{Z}_2$ .

Thus in this case the two semi-direct products  $\mathbb{Z}_5 \rtimes \mathbb{Z}_2 \times \mathbb{Z}_2$  corresponding to the trivial homomorphism case and  $\mathbb{Z}_5 \rtimes_1 (\mathbb{Z}_2 \times \mathbb{Z}_2)$  corresponding to the nontrivial homomorphism.

**Therefore there are five groups of order 20 upto isomorphism and they are as follows**

$$\mathbb{Z}_{20}, \mathbb{Z}_2 \times \mathbb{Z}_{10}, D_{10} \text{ and the semi direct products } \mathbb{Z}_5 \rtimes_1 \mathbb{Z}_4 \text{ and } \mathbb{Z}_5 \rtimes_2 \mathbb{Z}_4.$$

(d)

Let  $G$  be a group such that  $|G| = 30$ .

Then factorize the order of  $G$ ,

$$\begin{aligned} |G| &= 30 \\ &= 2 \times 3 \times 5 \end{aligned}$$

By Sylow Theorem,  $s_3 \equiv 1 \pmod{3}$  and  $s_3$  divides 10.

Thus the possible value for  $s_3$  is 1.

Similarly the possible value for  $s_5$  is 6, otherwise there would exist 20 elements of order 3 and 24 elements of order 5.

This implies that the number of elements in  $G$  is more than  $|G| = 30$ .

Hence a contradiction is obtained

So either the 3-Sylow subgroup  $P$  or the 5-Sylow subgroup  $Q$  is normal,

This implies that subgroup  $N = PQ$  of  $G$ .

Since  $\gcd(3, 5) = 1$ ,  $P \cap Q = \{1\}$  and therefore  $|N| = 15$

By Sylow theorem and since 5 is not congruent to 1 mod 3,  $N$  is cyclic.

Since its index in  $G$  is 2, it is a normal subgroup. Any 2-Sylow subgroup is a complement of  $N$  in  $G$ .

Let

$$\phi: \frac{\mathbb{Z}}{2\mathbb{Z}} \rightarrow \text{Aut}\left(\frac{\mathbb{Z}}{15\mathbb{Z}}\right) \text{ be a group homomorphism}$$

Then

$$G \cong \frac{\mathbb{Z}}{15\mathbb{Z}} \rtimes_{\phi} \frac{\mathbb{Z}}{2\mathbb{Z}}$$

Now since

$$\begin{aligned} \text{Aut}\left(\frac{\mathbb{Z}}{15\mathbb{Z}}\right) &\cong \left(\frac{\mathbb{Z}}{15\mathbb{Z}}\right)^{\times} \\ &\cong \left(\frac{\mathbb{Z}}{3\mathbb{Z}} \times \frac{\mathbb{Z}}{5\mathbb{Z}}\right)^{\times} \\ &\cong \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{4\mathbb{Z}} \end{aligned}$$

As  $\phi$  is a homomorphism and order of 1 in  $\mathbb{Z}/2\mathbb{Z}$  is 2, the order of its image under this homomorphism  $\phi(1)$  divides 2.

In  $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$  there exists four possible image given as follows

$$\begin{aligned} \phi(1) &= (0, 0) \\ \phi(1) &= (1, 0) \\ \phi(1) &= (1, 2) \\ \phi(1) &= (0, 2) \end{aligned}$$

Since  $\phi$  is completely determined by  $\phi(1)$ , hence there are at most 4 isomorphism types of groups of order 30.

Now consider the groups  $\mathbb{Z}_{30}$ ,  $D_{15}$ ,  $S_3 \times \mathbb{Z}_5$  and  $D_5 \times \mathbb{Z}_3$ ,

The order of every group mentioned above is 30.

Now since the number of elements of order 2 are different in every single one of them, in particular 1, 15, 3 and 5 are the number of elements of order 2 in the groups  $\mathbb{Z}_{30}$ ,  $D_{15}$ ,  $S_3 \times \mathbb{Z}_5$  and  $D_5 \times \mathbb{Z}_3$  respectively.

Since they all contain different number of elements of order 2, so they all are non-isomorphic.

**Therefore there are four groups of order 30 upto isomorphism and they are as follows**

$\mathbb{Z}_{30}$ ,  $D_{15}$ ,  $S_3 \times \mathbb{Z}_5$  and  $D_5 \times \mathbb{Z}_3$ .

10. a

A group  $G$  is said to be simple only when its only normal subgroups are the trivial group and the group itself.

Mathematically,  $G$  is a simple group if  $\{e\}, G$  are the only normal subgroups of  $G$ .

[Comment](#)

## Step 2 of 7 ^

Let  $G$  be a group such that  $|G| < 60$  and  $|G|$  is a non-prime number.

Clearly for all such groups the order of the group is divisible by a prime. Consider the non-prime numbers less than 60, starting with larger prime divisor

For  $29, |G| = 29 \times 2 = 58$ .

Since index of the  $29$ -sylow subgroup is the smallest prime divisor of  $|G|$ , thus  $29$ -sylow subgroup is normal.

For  $23, |G| = 23 \times 2 = 46$ .

On similar account of reasoning as for  $29$ , the  $23$ -sylow subgroup is normal.

For  $19$  two possibilities exists

First for  $19, |G| = 19 \times 2 = 38$  and secondly  $19, |G| = 19 \times 3 = 57$

In both the cases  $19$ -sylow subgroup is normal

Similarly  $17$ -sylow subgroup is normal.

For  $13$  there are three possibilities

First for  $13, |G| = 13 \times 2 = 26$  and secondly  $13, |G| = 13 \times 3 = 39$

In both the cases  $13$ -sylow subgroup is normal

Thirdly  $13, |G| = 13 \times 4 = 52$ , the number of  $13$ -sylow subgroup  $s_{13}$  divides  $4$  and  $s_{13} \equiv 1 \pmod{13}$

Hence there exists only one  $13$ -sylow subgroup and so it is normal.

Similarly for  $7, 11$  the  $7$ -sylow subgroup and  $11$ -sylow subgroup is normal.

Thus generalize above notion : If order of the group is a product  $pn$  where  $n < p$ , then the  $p$ -sylow subgroup is normal.

Now if  $|G| = 7^2$

Then a subgroup of order  $7$  is normal since its index,  $7$ , is the smallest prime divisor of group  $G$ .

Hence if the order of the group is of the form  $|G| = p^m$ , where  $p$  is a prime and  $m \geq 2$



By Cauchy's theorem

$G$  has a subgroup of order  $p^{m-1}$ , which is normal in  $G$  as its index  $p$  in  $G$  is the smallest prime divisor of  $G$ .

If  $|G| = 7(8) = 56$

Then there is either 1 or 8 sylow subgroups.

If there is only 1, it is normal.

If there are 8, there are  $8 \times 6 = 48$  elements of order 7. Hence the remaining 8 elements form a subgroup of order 8.

Hence, those elements form a normal subgroup since the elements not of order 8 are transformed among themselves by conjugation, which preserves order.

Thus, the 2-Sylow subgroup will be normal.

For 5, in the groups which have order of the form  $5n$ , where  $n = 2, 3, 4, 8, 9$ , there is only 1 5-sylow subgroup and thus it is normal.

If  $|G| = 5^2 = 25$ ,  $G$  is not simple.

If  $|G| = 5(6) = 30$ ,  $G$  has 1 or 6 5-sylow subgroups. If  $G$  has 6 5-sylow subgroups then it has 24 elements of order 5, which leaves 6 elements that are permuted among themselves in terms of conjugation.

By Cauchy's Theorem there is a subgroup of order 3, which is normal

Since if this doesn't hold then there would be at least 4 3-sylow subgroups which would give at least 8 elements of order 3, while only 6 elements are there.

If  $|G| = 5(10) = 5^2 \cdot 2 = 50$ , the 5-sylow subgroup is normal since its index is 2.

For 3,

$|G| = 3^n 2^m$ , with  $n \geq 1$

If  $m = 0, n \geq 2$ , the group is not simple since the order is a power of a prime.

If  $m = 1$ , the 3-sylow subgroup is normal, since its index is 2

If  $m = 2$  and  $n = 1$ , there is either 1 or 4 3-sylow subgroups.

Now if there is 1 3-sylow subgroup, it is normal.

If there are 4 3-sylow subgroups, there are 8 elements of order 3, and the remaining 4 elements must form a subgroup of order 4, which is unique

If  $m = 2$  and  $n = 2$ , then order of the group is 36.

Thus there are either 1 or 4 3-sylow subgroups.

If there are 4 3-sylow subgroups, then the stabilizer subgroup  $C$  of a 3-sylow subgroup under conjugation has 9 elements.

Consider a homomorphism  $\phi: G \rightarrow \text{Perm}(G/C)$ .

Since  $|G| = 36$  and  $|\text{Perm}(G/C)| = 24$ , the kernel of this homomorphism has more than one element and also  $\ker(\phi) \subset C$ , it is not the whole group.

Thus the kernel is a normal subgroup of  $G$  which implies that  $G$  is not simple.

If  $m = 3$  and  $n = 1$ , then order of the group is 24, then there are either 1 or 4 3-sylow subgroups.

If there are 4 3-sylow subgroups, then the stabilizer subgroup  $C$  of a 3-sylow subgroup under conjugation has 6 elements.

Consider a homomorphism  $\phi: G \rightarrow \text{Perm}(G/C)$ .

Since  $|G| = 24$  and  $|\text{Perm}(G/C)| = 24$

Now if the map is bijective then  $G$  is  $S_4$  which is not simple as it contains the normal subgroup  $A_4$ .

If the map is not bijective, the kernel of this homomorphism has more than one element and also  $\ker(\phi) \subset C$ , it is not the whole group.

Thus the kernel is a normal subgroup of  $G$  which implies that  $G$  is not simple.

If  $m = 4$  and  $n = 1$ , then order of the group is 48.

So there are either 1, 4, 16 3-sylow subgroups.

If there are 16 3-sylow subgroups then  $G$  has  $16(2) = 32$  element of order 3, which leaves 16 elements to form a 2-sylow subgroup.

Hence, there is room for only one 2-sylow subgroup, which is then normal.

If there are 4 3-sylow subgroups, then the stabilizer subgroup  $C$  has 12 elements and  $G/C$  has 4 elements.

Again as in the case of group of order 36 this leads to a normal subgroup of  $G$ .

Finally for 2,  $|G| = 2^n, n \geq 2$

These groups are not simple since order of these groups is a power of a prime.

**Therefore, the only simple groups of order  $< 60$  are the groups of prime order.**

## Section 8

1. a

Consider the groups of order 12.

[Comment](#)

Step 2 of 2 ^

Since,

The dihedral group,

$$\begin{aligned} D_6 &= \{x, y, z : x^2 = y^3 = z^2 = 1, yz = zy, xz = zx, xy = y^2x\} \\ &\cong \{x, y : x^2 = y^3 = 1, xy = y^2x\} \times \{z : z^2 = 1\} \\ &\cong S_3 \times C_2 \end{aligned}$$

Therefore,  $D_6 \cong S_3 \times C_2$ .

2. a

!!!

3. a

Consider the class equations of the groups of order 12.

The possible class equation of order 12:

Here,

First possibility:

$$1+1+1+1+1+1+1+1+1+1+1+1$$

This will be the class equation of cyclic group  $C_{12}$ , product of two cyclic groups  $C_2 \times C_6$  and abelian group of order 12.

As,

The number of conjugacy classes of size one is 12.

So,

The total number of conjugacy classes is 12.

Now,

Second possibility:

$$1+1+2+2+3+3$$

This will be the class equation of dicyclic group of order 12 and dihedral group  $D_{12}$ .

Since,

The number of conjugacy classes of size one is 2, number of conjugacy classes of size 2 is 2 and number of conjugacy classes of size 3 is 2.

Thus,

The total number of conjugacy classes is 6.

Next,

Third possibility:

$$1+3+4+4$$

This will be the class equation of alternating group  $A_4$ .

Then,

The number of conjugacy classes of size one is 1, the number of conjugacy classes of size 3 is 1 and the number of conjugacy classes of size 4 is 2.

Therefore,

The total number of conjugacy classes is 4.

Hence, the possible class equations of group of order 12 are  $1+1+2+2+3+3$

$$1+1+1+1+1+1+1+1+1+1+1+1 \text{ and } 1+3+4+4.$$

4. a

Consider the every group of order  $2p$ .

Since,

The order of group  $|K| = 2p$

From Cauchy's theorem  $K$  will have an element  $a \in K$  of order 2 ,

And,

An element  $b \in K$  of order  $p$  ,

As,  $|\langle a \rangle \cap \langle b \rangle|$  will divide 2 and  $p$  ,

So,

$$|\langle a \rangle \cap \langle b \rangle| = 1$$

Then,

$$\begin{aligned} |\langle a \rangle \langle b \rangle| &= \frac{|\langle a \rangle| |\langle b \rangle|}{|\langle a \rangle \cap \langle b \rangle|} \\ &= 2p \end{aligned}$$

Thus,

$$G = |\langle a \rangle \langle b \rangle|$$

Now,  $\langle b \rangle$  is normal.

Also, for some  $k \in \mathbb{Z}$

$$aba^{-1} = b^k$$

Thus,

$$\begin{aligned} b^{k^2} &= (b^k)^k \\ &= (aba)(aba) \dots (aba) \\ &= aba^2ba^2b \dots a^2ba \\ &= abb \dots ba \\ &= ab^k a \\ &= a(aba)a \\ &= a^2ba^2 \\ &= ebe \\ &= b \end{aligned}$$

Further,

$$b^{k^2-1} = e$$

But,

$$|\langle b \rangle| = p$$

As a result,  $p$  divides  $k^2 - 1$ .

And,  $k^2 - 1 = (k-1)(k+1)$

Due to prime number,  $p$  will divide either  $k-1$  or  $k+1$ .

If  $p$  divides  $k+1$ ,

Then,

$$\begin{aligned} aba &= b^k \\ &= b^{k+1-1} \\ &= b^{k+1}b^{-1} \\ &= b^{-1} \end{aligned}$$

Thus,

$$\begin{aligned} K &= \{a, b : a^2 = b^p = e, aba^{-1} = b^{-1}\} \\ &\cong D_{2p} \end{aligned}$$

So,  $K$  is dihedral.

If  $p$  divides  $k-1$ ,

Then,

$$\begin{aligned} aba &= b^k \\ &= b^{k-1+1} \\ &= b^{k-1}b \\ &= b \end{aligned}$$

After multiplying  $a^{-1}$  on both sides,

$$\begin{aligned} ab &= ba^{-1} \\ &= ba \end{aligned}$$

Therefore, the group  $K$  is abelian.

But,

$$G = \langle ab \rangle = 2p$$

So,  $K$  is cyclic.

Hence, a group of order  $n = 2p$ , where  $p$  is prime, is either cyclic or dihedral.

5. a

!!!

6. a

Consider  $G$  be a group of order 55.

(a)

Since,

$$O(G) = 55 = 5 \cdot 11$$

By the third Sylow theorem,

The number of Sylow 11-subgroup  $\langle 11 \rangle$  is 1.

And,

The number of Sylow 5-subgroup  $\langle 5 \rangle$  is either 1 or 5.

Since,

$$O(\langle 5 \rangle \cap \langle 11 \rangle) = 1$$

As,

$$O(\langle 5 \rangle \langle 11 \rangle) = \frac{O(\langle 5 \rangle) O(\langle 11 \rangle)}{O(\langle 5 \rangle \cap \langle 11 \rangle)}$$

$$O(\langle 5 \rangle \langle 11 \rangle) = \frac{O(\langle 5 \rangle) O(\langle 11 \rangle)}{1}$$

So,

$$\begin{aligned} O(\langle 5 \rangle \langle 11 \rangle) &= O(\langle 5 \rangle) O(\langle 11 \rangle) \\ &= 55 \end{aligned}$$

Therefore,

$$G \approx \langle 5 \rangle \times \langle 11 \rangle$$

As,

The subgroup  $\langle 11 \rangle$  is normal subgroup of  $G$ .

Let,

$$x \in \langle 11 \rangle, y \in \langle 5 \rangle$$

Then,

$$y \langle 11 \rangle = \langle 11 \rangle y$$

$$y \langle 11 \rangle y^{-1} = \langle 11 \rangle$$

Thus,

$$yxy^{-1} = x^r, 1 \leq r < 11$$

Hence,  $G$  is generated by two elements  $x$  and  $y$  with  $x^{11} = 1, y^5 = 1, yxy^{-1} = x^r$  for some  $r$ ,

$$1 \leq r < 11.$$

(b)

Since,

$$\begin{aligned} x &= y^5 xy^{-5} \\ &= y^4 (yxy^{-1}) y^{-4} \\ &= y^4 x^r y^{-4} \\ &= y^3 (yxy^{-1}) y^{-3} \end{aligned}$$

Now,

$$\begin{aligned} x &= y^3 x^{r^2} y^{-3} \\ &= y^2 (yxy^{-1}) y^{-2} \\ &= y^2 x^{r^3} y^{-2} \\ &= y (yxy^{-1}) y^{-1} \end{aligned}$$

And,

$$\begin{aligned} x &= yx^{r^4} y^{-1} \\ &= x^{r^5} \end{aligned}$$

Thus,

$$r^5 \equiv 1 \pmod{11}$$

Therefore, the values of  $r$ , for which  $r^5 \not\equiv 1 \pmod{11}$  are not possible.



(c)

Since,

There are two isomorphism classes of groups of order 55:

$$G \approx \langle 5 \rangle \times \langle 11 \rangle$$

And,

$$G \approx \langle x^{11} = 1, y^5 = 1, yxy^{-1} = x^r, r^5 = 1 \pmod{11} \rangle$$

Hence, there are two isomorphism classes of groups of order 55.

## Section 9

1. a

!!!

2. a

Let  $S = \{a, b, c, \dots\}$  be a basis for the free group  $F(S)$ . The set  $F(S)$  of the equivalence classes of words in the set  $S' = \{a, a^{-1}, b, b^{-1}, c, c^{-1}, \dots\}$  is a group, with the law of composition induced from juxtaposition in  $S' = \{a, a^{-1}, b, b^{-1}, c, c^{-1}, \dots\}$ .

Let  $S = \{a, b, c, \dots\}$  denotes the set of generators for the free group.

Let  $P$  denote the set of all reduced closed words in  $S$ .

Let  $Q$  denote the set of conjugacy classes of words in  $S$ .

Define a map  $\phi: Q \rightarrow P$  by joining the ends of any element in a given conjugacy class and then reduce it.

Well-Defined

Let  $x, y$  belong to same conjugacy class.

This implies that  $zxz^{-1} = y$  for some word  $z \in S$ .

Close  $y$  yields the word  $y$  only. Now close the word  $zxz^{-1}$  and reduce it, this yields the word  $x$  only. Hence they both have the same reduced form.

Thus the map is well-defined.

Injection

Let  $x, y \in S$  be two reduced words such that  $\phi(x) = \phi(y)$ , where  $x, y$  are to be treated as representative elements for two conjugacy classes.

Now the only possible cancellation that can occur when  $x, y$  are closed is of the form  $x = w^{-1}zw$  being sent to the closed word formed from  $z$ .

Thus  $x, y$  are conjugates.

Hence they belong to the same conjugacy classes.

Thus the map defined above is injective.

Surjective

Given any closed reduced word, break it to an arbitrary point and consider the conjugacy classes of the word that arises.

This conjugacy class will clearly map to the given reduced closed word.

Hence the map defined above is surjective also.

Thus  $\phi: Q \rightarrow P$  is bijective in nature.

**Therefore, there exists a bijective correspondence  $\phi$  between reduced closed words and the conjugacy classes in the free group given by joining the ends of any element in a given conjugacy class and then reduce it.**

## Section 10

1. a

Consider, the group  $G = \langle x, y, xyx^{-1}y^{-1} \rangle$  is called a free abelian group and  $u, v$  elements of abelian group  $A$ .

[Comment](#)

Step 2 of 2 ^

Suppose there is not unique homomorphism. Let there are two homomorphism  $\phi_1$  and  $\phi_2$ .

And,

$$\begin{aligned}\phi(xy x^{-1} y^{-1}) &= \phi_1(x) \phi_1(y) \phi_1(x^{-1}) \phi_1(y^{-1}) \\ &= \phi_1(x) \phi_1(y) [\phi_1(x)]^{-1} [\phi_1(y)]^{-1} \\ &= uu^{-1}vv^{-1} \\ &= 1\end{aligned}$$

So, the element  $xyx^{-1}y^{-1}$  maps to identity element of  $A$ .

And,

As,  $G$  be a free abelian group therefore  $xyx^{-1}y^{-1}$  will generate identity element of  $G$  Therefore, the group  $\text{Ker}\phi$  will have elements generated by  $xyx^{-1}y^{-1}$ ,

Hence, **there is a unique homomorphism  $\phi: G \rightarrow A$  by mapping property of free abelian group.**

2. a

Consider  $\varphi: G \rightarrow G'$  be a surjective group homomorphism and  $S$  be a subset of  $G$  whose image  $\varphi(S)$  generates  $G'$  and let  $T$  be a set of generators of  $\ker \varphi$ .

[Comment](#)

Step 2 of 2 ^

Since,

The mapping  $\varphi: G \rightarrow G'$  is a surjective group homomorphism.

And,

The set  $S$  be a subset of  $G$  whose image  $\varphi(S)$  generates  $G'$ .

As,

$$\dim G' + \dim(\ker \varphi) = \dim G$$

Also,

Suppose  $T$  be a set of generators of  $\ker \varphi$ .

Therefore,

$$|S \cup T| = \dim G$$

Hence, **the set  $S \cup T$  generates  $G$ .**

3. a

Consider every finite group  $G$  representation.

[Comment](#)

Step 2 of 2 ^

Suppose,

The finite group  $G$  is generated by set  $K$ .

And,

The set of relations between elements in  $K$  is  $P$ .

Now,

A presentation is called finitely generated if  $K$  is finite and finitely generated if  $P$  is finite.

As,

The set  $K$  will always be finite for a finite group  $G$ ,

But,

There may be infinitely relations between elements in  $K$ .

So,

The set  $P$  may be finite or infinite.

Therefore, **every finite group  $G$  will be presented by a finite set of generators but can be represented by infinite set of relations.**

4. a

### Mapping property of the free groups

Let  $F$  be the free group on a set  $S = \{a, b, \dots\}$ , and let  $G$  be a group.

Any map  $f: S \rightarrow G$  extends in a unique way to a group homomorphism  $\varphi: F \rightarrow G$ .

Here the meaning of the map can be understood in the way if  $f(x) = x_o$  denotes the image of an element  $x$  of the set  $S$  then  $\varphi$  sends the word in  $S' = \{a, a^{-1}, b, b^{-1}, \dots\}$  to the corresponding product of elements  $\{a_o, a_o^{-1}, b_o, b_o^{-1}, \dots\}$  in the group  $G$ .

Let  $G = \langle x, y; xyx^{-1}y^{-1} \rangle$  be a free abelian group.

Let  $A$  be an abelian group

Let  $u, v \in A$

Define a map  $f: \{x, y\} \rightarrow A$  as follows

$$f(x) = u$$

$$f(y) = v$$

Then this map can be extended to a unique homomorphism  $\varphi: G \rightarrow A$  from the above mentioned mapping property of free groups.

Also since  $f$  maps the letter  $x$  to  $u$  and the letter  $y$  to  $v$

By the definition of the extension of the map  $f: \{x, y\} \rightarrow A$  to the homomorphism  $\varphi: G \rightarrow A$  the following conditions hold

$$\varphi(x) = u$$

$$\varphi(y) = v$$

Hence if  $u, v$  are elements of an abelian group then there exists a unique homomorphism

$$\varphi: G \rightarrow A \text{ such that } \varphi(x) = u \text{ and } \varphi(y) = v$$

**Therefore, the result in the question has been proved.**

5. a

The group generated by  $x, y, z$  with the single relation  $yx y z^{-2} = 1$ .

[Comment](#)

Step 2 of 2 ^

If  $x = y^{-1}z^2y^{-1}$  then,

$$\begin{aligned} yx y z^{-2} &= y y^{-1} z^2 y^{-1} y z^{-2} \\ &= z^2 z^{-2} \\ &= 1 \end{aligned}$$

So, the group is generated by  $y$  and  $z$ .

Therefore, **the group generated by  $x, y, z$  with the single relation  $yx y z^{-2} = 1$  is actually a free group generated by  $y, z$ .**

6. a

Consider a subgroup  $H$  of a group  $G$  is characteristic if it is carried to itself by all automorphisms of  $G$ .

(a)

Assume a subgroup  $H$  of a group  $G$  is characteristic subgroup.

Let  $f$  be an automorphism of  $G$ .

And  $h_a$  is conjugation by  $a$ .

Then,

By definition of automorphism and conjugation,

$$fh_a f^{-1}(g) = h_{f(a)}(g)$$

Since,

A subgroup  $H$  of a group  $G$  is characteristic if it is carried to itself by all automorphisms of  $G$ .

So,

For  $g \in G$ ,

$$gHg^{-1} = H$$

Therefore, every characteristic subgroup is normal.

And,

Consider  $\phi$  be an automorphism.

Then, it will have to prove that,

$$\phi(Z(G)) = Z(G)$$

By the commutative property of center  $Z(G)$ ,

$$\phi(Z(G)) \subseteq Z(G)$$

Also, center  $Z(G)$  is normal subgroup.

As a result, for  $g \in G$ ,

$$gZg^{-1} = Z$$

Now,

By definition of automorphism and conjugation,

$$fh_a f^{-1}(g) = h_{f(a)}(g)$$

Thus,

$$Z(G) \subseteq \phi(Z(G))$$

Therefore,

$$\phi(Z(G)) = Z(G)$$

Hence, center  $Z$  is a characteristic subgroup.

(b)

The quaternion group:

$$\{1, -1, i, -i, j, -j, k, -k\}$$

The normal subgroups of quaternion group are:

$$\{1, -1, i, -i\}$$

$$\{1, -1, j, -j\}$$

$$\{1, -1, k, -k\}$$

Because, for these three groups:

If  $g \in G$  then,

$$gHg^{-1} = H$$

And,

The characteristic subgroup of quaternion group is center  $Z$  :

$$\{1, -1\}$$

Because,

The center  $Z$  is characteristic subgroup.

As,

It is carried to itself by all automorphisms of  $G$ .

Hence, **there are three normal subgroups and one characteristic subgroup of quaternion group.**

7. a

Let  $G$  be a group and  $x, y \in G$ . The commutator of  $x, y$  is denoted by  $[x, y]$  and is defined by

$$[x, y] = xyx^{-1}y^{-1}.$$

The Commutator subgroup  $C$  of a group  $G$  is the subgroup of  $G$  generated by the commutators  $\{[x, y] | x, y \in G\}$ . Also it is the smallest subgroup of  $G$  containing all the commutators.

Let  $G$  be a group.

Define  $A = \{xyx^{-1}y^{-1} | x, y \in G\}$

Now let  $T : G \rightarrow G$  be an arbitrary automorphism and hence an isomorphism.

Let  $g \in A$  be arbitrary then  $g$  is of the form  $g = xyx^{-1}y^{-1}$  for some  $x, y \in G$

Now consider  $T(g)$

$$\begin{aligned} T(g) &= T(xyx^{-1}y^{-1}) \\ &= T(x)T(y)T(x^{-1})T(y^{-1}) \\ &= T(x)T(y)(T(x))^{-1}(T(y))^{-1} \end{aligned}$$

Since  $T(x), T(y) \in G$  and  $T(g) = pqp^{-1}q^{-1}$  where  $p = T(x), q = T(y)$  and thus lie in  $G$ .

Hence  $T(g) \in A$  for all  $g \in A$

This implies that  $T(A) \subset A$

Now to show  $A \subset T(A)$  is equivalent to show  $T^{-1}(A) \subset A$ .

Now since  $T : G \rightarrow G$  is an automorphism so  $K = T^{-1} : G \rightarrow G$  is also an automorphism.

Use the fact proved above  $K(A) \subset A$

Hence  $T^{-1}(A) \subset A$  this implies  $A \subset T(A)$ .

Thus  $T(A) = A$

Since  $C$  is the smallest subgroup of  $G$  containing all the commutators.

So  $T(C) = C$

Now since every characteristic subgroup is normal, so  $C$  is normal.



Now let  $xC, yC \in G/C$  be arbitrary elements.

Since  $xyx^{-1}y^{-1} \in C$  so

$$xyx^{-1}y^{-1}C = C$$

Pre-multiply on both sides by  $x^{-1}$

$$x^{-1}xyx^{-1}y^{-1}C = x^{-1}C$$

$$(x^{-1}x)yx^{-1}y^{-1}C = x^{-1}C$$

$$yx^{-1}y^{-1}C = x^{-1}C$$

Pre-multiply on both sides by  $y^{-1}$

$$y^{-1}yx^{-1}y^{-1}C = y^{-1}x^{-1}C$$

$$(y^{-1}y)x^{-1}y^{-1}C = y^{-1}x^{-1}C$$

$$x^{-1}y^{-1}C = y^{-1}x^{-1}C$$

Thus  $x^{-1}y^{-1}C = y^{-1}x^{-1}C$  implies that  $(yx)^{-1}C = (xy)^{-1}C$

Also  $C$  is normal, so  $xyC = yxC$

Hence  $xC, yC \in G/C$  commute with each other.

Thus  $G/C$  is abelian.

**Therefore, the commutator subgroup  $C$  of a group  $G$  is a characteristic subgroup and  $G/C$  is an abelian group.**

8. a

!!!

9. a

!!!

10. a

!!!

## Section 11

1. a

Todd coexter algorithm-Let  $G$  be a finite group with generators and relation. Let  $H$  be cyclic subgroup of  $G$ . Then following rules must be followed while using Todd Coexter Algorithm

- 1) The operation of each generator is a permutation.
- 2) The relations operate trivially: they fix every coset.
- 3) The generators of  $H$  fix coset  $[H]$ .
- 4) The operation is transitive.

Define a group  $G$  given by

$$G = \langle x, y \mid x^3, y^3, yxyxy \rangle$$

Let  $H = \langle y \rangle$

Now let the co-set  $[H]$  be denoted by  $\mathbf{1}$ .

Now in tabular form

Let  $\mathbf{1}, \mathbf{2}, \mathbf{3}$  denote the various distinct indices of the co-sets of  $G$ .

$y \quad y \quad y$

---



---

$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$
$\mathbf{2}$	$\mathbf{3} = \mathbf{2}$	$\mathbf{2}$	$\mathbf{2}$

---



---

$x \quad x \quad x$

$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3} = \mathbf{2}$	$\mathbf{1}$
$\mathbf{2}$	$\mathbf{3} = \mathbf{2}$	$\mathbf{1}$	$\mathbf{2}$

---



---

$y \quad x \quad y \quad x \quad y$

$\mathbf{1}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$	$\mathbf{1}$	$\mathbf{1}$
$\mathbf{2}$	$\mathbf{3}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$

Since  $yxyxy = 1$ ,

Thus  $\mathbf{3} = \mathbf{2}$

Also since  $x$  sends  $\mathbf{2}$  to  $\mathbf{1}$ , thus  $\mathbf{3} = \mathbf{2} = \mathbf{1}$ .

Hence there exists only one co-set  $[H]$  of  $G$ .

This implies that  $H = \langle y \rangle = G$

Since  $O(y) = 3$ , thus  $G$  is a cyclic group of order 3.

**Therefore,  $G = \langle x, y \mid x^3, y^3, yxyxy \rangle$  is a cyclic group of order 3.**

## 2. a

Let  $D_n$  be a dihedral group of order  $2n$ . Then  $D_n$  is generated by two elements say  $x$  and  $y$  that satisfy the following conditions

$$x^n = 1, y^2 = 1 \text{ and } yx = x^{-1}y$$

$$\text{Thus } D_n = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$$

Double Coset of a group with respect two subgroups  $H, K$  is denoted by  $HgK$  where  $g$  is an arbitrary element of the group and is defined as follows

$$HgK = \{h g k : h \in H, k \in K\}$$

Let  $D_n$  be a dihedral group. Let  $H = \{1, y\}$  be a subgroup of  $D_n$ .

The possible double coset  $HgH$  are given as follows

$\{g, yg, gy, ygy\}$  where  $g \in D_n$  denotes an arbitrary element

**Case 1**

For  $HgH = g$

Clearly the whole double coset is equal to the space  $D_n$ .

**Case 2**

For  $HgH = yg$

Let  $g \in D_n$  be an arbitrary element then  $g \in \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

If  $g = x^i$ ,  $0 \leq i \leq n-1$  then

$$\begin{aligned} yg &= yx^i \\ &= x^{-i}y \end{aligned}$$

If  $g = y$  then

$$\begin{aligned} yg &= y^2 \\ &= 1 \end{aligned}$$

If  $g = x^i y$ ,  $0 \leq i \leq n-1$

$$\begin{aligned} yg &= yx^i y \\ &= yx(x)^{i-1} y \\ &= x^{-1}y(x)^{i-1} y \end{aligned}$$

Continuous simplification as above implies that  $yg = x^i$

Thus  $HgH = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\} = D_n$

**Case 3**

For  $HgH = gy$

This case is same as the second one.

Hence,  $HgH = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\} = D_n$

**Case 4**

For  $HgH = ygy$

Let  $g \in D_n$  be an arbitrary element then  $g \in \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

If  $g = x^i$ ,  $0 \leq i \leq n-1$  then

$$\begin{aligned} ygy &= yx^i y \\ &= yx(x)^{i-1} y \\ &= x^{-1}y(x)^{i-1} y \end{aligned}$$

Continuous simplification as above implies that  $ygy = x^i$

If  $g = y$  then  $ygy = y^3 = 1$

If  $g = x^i y$ ,  $0 \leq i \leq n-1$  then,

$$\begin{aligned} y(x^i y) &= yx^i y^2 \\ &= yx^i \\ &= x^{-i}y \end{aligned}$$

Thus in this case also

$$HgH = D_n$$

Now use the formula  $|HgK| = |H| \cdot [K : K \cap g^{-1}Hg]$

Since here  $H = K$ ,

$$\text{So } |HgH| = |H| \cdot [H : H \cap g^{-1}Hg]$$

Now

$$[H : H \cap g^{-1}Hg] = \frac{|H|}{|H \cap g^{-1}Hg|}$$

Since  $|H| = 2$  and  $|H \cap g^{-1}Hg| = \{1, 2\}$

$$\text{Hence } [H : H \cap g^{-1}Hg] = \{1, 2\}$$

$$\text{Thus } |HgH| = \{2, 4\}$$

Therefore the double coset of the subgroup  $H = \{1, y\}$  of the dihedral groups  $D_n$  are the group itself and the number of elements in double coset of the subgroup  $H = \{1, y\}$  is either 2 or 4.

### 3. a

Todd coexter algorithm-Let  $G$  be a finite group with generators and relation. Let  $H$  be cyclic subgroup of  $G$ . Then following rules must be followed while using Todd Coexter Algorithm

- 1) The operation of each generator is a permutation.
- 2) The relations operate trivially: they fix every coset.
- 3) The generators of  $H$  fix coset  $[H]$ .
- 4) The operation is transitive.

(a)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^2 = y^2, xyx = yxy \rangle$$

Here  $xyx = yxy$  can be re-written as  $xyxy^{-1}x^{-1}y^{-1} = 1$

Use Todd-Coexter Algorithm

Let  $H = \langle x \rangle$  be the chosen subgroup.

Let  $\mathbf{1}$  denote the index corresponding to the co-set which  $x$  fixes.

Now draw the table as follow

$x \quad x$

---



---

1	1	1
2	3	2
3	2	3

$y \ y$

---



---

1	2	1
2	1	2
3	3	3

---



---

$x$	$y$	$x$	$y^{-1}$	$x^{-1}$	$y^{-1}$
1	1	2	3	3	2
2	3	3	2	1	1
3	2	1	1	2	3

Thus the Todd-Coexter algorithm terminates after three cosets

Also, in cyclic notation  $x = (23)$  and  $y = (12)$ , so order of  $x = 2$  as well as order of  $y = 2$

Clearly  $[G : H] = 3$  since there are three index namely **1, 2, 3**

$$\begin{aligned}
 |G| &= [G : H]|H| \\
 &= 3 \times 2 \\
 &= 6
 \end{aligned}$$

**Therefore, the order of the above defined group is 6 .**

**(b)**

Let  $G$  be a group given by

$$G = \langle x, y \mid x^3 = y^2, xyx = yxy \rangle$$

Here  $xyx = yxy$  can be re-written as  $xyxy^{-1}x^{-1}y^{-1} = 1$

Use Todd-Coexter Algorithm

Let  $H = \langle y \rangle$  be the chosen subgroup.

Let **1** denote the index corresponding to the co-set which  $y$  fixes.

Now draw the table as follow

*x x x*

---

---

1	2	3	1
2	3	1	2
3	1	2	3
4	4	4	4
5	7	6	5
6	5	7	6
7	6	5	7
8	8	8	8

*y y*

---

---

1	1	1	1
2	4	5	2
4	5	2	4
5	2	4	5
3	6	8	3
6	8	3	6
7	7	7	7
8	3	6	8



---



---

$x \quad y \quad x \quad y^{-1} \quad x^{-1} \quad y^{-1}$

1	2	4	4	2	1	1
2	3	6	5	4	4	2
3	1	1	2	2	6	3
4	4	5	7	7	5	4
5	7	7	6	3	2	5
6	5	2	3	8	8	6
7	6	8	8	6	7	7
8	8	3	1	1	3	8

Thus the Todd-Coexter algorithm terminates after eight cosets

Also, in cyclic notation  $x = (123)(567)$  and  $y = (245)(368)$ , so order of  $x = 3$  as well as order of  $y = 3$

Clearly  $[G : H] = 8$  since there are eight index namely **1, 2, 3, 4, 5, 6, 7, 8**

$$\begin{aligned}
 |G| &= [G : H]|H| \\
 &= 8 \times 3 \\
 &= 24
 \end{aligned}$$

Therefore, the order of the above defined group is **24**.

(c)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^4 = y^2, xyx = yxy \rangle$$

Here  $xyx = yxy$  can be re-written as  $xyxy^{-1}x^{-1}y^{-1} = 1$

Use Todd-Coexter Algorithm

Let  $H = \langle x \rangle$  be the chosen subgroup.

Let **1** denote the index corresponding to the co-set which  $x$  fixes.

Now draw the table as follow

$x \quad x \quad x \quad x$

---



---

1	1	1	1	1
2	3	2	3	2
3	2	3	2	3

$y^{-1} y$

---



---

1	2	1
2	1	2
3	3	3

---



---

$x \quad y \quad x \quad y^{-1} \quad x^{-1} \quad y^{-1}$

1	1	2	3	3	2	1
2	3	3	2	1	1	2
3	2	1	1	2	3	3

Thus the Todd-Coexter algorithm terminates after eight cosets

Also, in cyclic notation  $x = (23)$  and  $y = (12)$ , so order of  $x = 2$  as well as order of  $y = 2$

Clearly  $[G:H] = 3$  since there are eight index namely **1,2,3**

$$\begin{aligned}
 |G| &= [G:H]|H| \\
 &= 2 \times 3 \\
 &= 6
 \end{aligned}$$

Therefore, the order of the above defined group is **6** .

(d)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^4 = y^4 = x^2 y^2 = 1 \rangle$$

Use Todd-Coexter Algorithm

Let  $H = \langle x \rangle$  be the chosen subgroup.

Let **1** denote the index corresponding to the co-set which  $x$  fixes.

Now draw the table as follow

$x \quad x \quad x \quad x$

---



---

1	1	1	1	1
2	3	2	3	2
3	2	3	2	3
4	5	4	5	4
5	4	5	4	5
6	7	6	7	6

$y \ y \ y \ y$

---



---

1	2	1	2	1
2	1	2	1	2
3	4	3	4	3
4	3	4	3	4
5	6	5	6	5
6				6

$x \ x \ y \ y$

---



---

1	1	1	2	1
2	3	2	1	2
3	2	3	4	3
4	5	4	3	4
5	4	5	6	5
6	7	6	5	6

This process does not terminate since different index keeps generating in the above algorithm.

Let in general the number of index be  $n$  namely **1,2,3,4,5,6,7,8,...,n**

Also, in cyclic notation  $x = (23)(45)(67) \dots ((n-1)n)$  and  $y = (12)(34)(56) \dots ((n-1)n)$ , so order of  $x = 2$  as well as order of  $y = 2$

Clearly  $[G : H] = n$  since there are  $n$  index namely **1,2,3,4,5,6,7,8,...,n**

$$\begin{aligned}
 |G| &= [G : H] |H| \\
 &= n \times 2 \\
 &= 2n
 \end{aligned}$$

Therefore, the order of the above defined group is  $2n$ .

(e)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^3 = y^2 = yxyxy = 1 \rangle$$

Use Todd-Coexter Algorithm

Let  $H = \langle y \rangle$  be the chosen subgroup.

Let  $\mathbf{1}$  denote the index corresponding to the co-set which  $y$  fixes.

Now draw the table as follow

$x \quad x \quad x$

---

---

1	2	3	1
2	3	1	2
3	1	2	3

$y \quad y$

---

---

1	1	1
2	3	2
3	2	3

$y \quad x \quad y \quad x \quad y$

---

---

1	1	2	3=2	1=2=3	1
2	3	1	1	2	3=2
3	2	1	1	2	3

Here,  $\mathbf{1} = \mathbf{2} = \mathbf{3}$

Thus the Todd-Coexter algorithm terminates after three cosets which are same

Also, in cyclic notation  $x = (\mathbf{1})$  and  $y = (\mathbf{1})$ , so order of  $x = 1$  as well as order of  $y = 1$

Clearly  $[G : H] = 1$  since there are eight index namely  $\mathbf{1}$

$$\begin{aligned} |G| &= [G : H] |H| \\ &= 1 \times 1 \\ &= 1 \end{aligned}$$

Therefore, the order of the above defined group is  $\mathbf{1}$ .

(f)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^3 = y^3 = yxyxy = 1 \rangle$$

Use Todd-Coexter Algorithm

Let  $H = \langle y \rangle$  be the chosen subgroup.

Let  $\mathbf{1}$  denote the index corresponding to the co-set which  $y$  fixes.

Now draw the table as follow

$x \quad x \quad x$

---



---

1	2	3	1
2	3	1	2
3	1	2	3
4			4

$y \quad y \quad y$

---



---

1	1	1	1
2	3	4	2
3	4	2	3
4	2	3	4

$y \quad x \quad y \quad x \quad y$

---



---

1	1	2	3=2	1=2=3=4	1
2	3	1	1	2	3=2
3	2	1	1	2	3
4	3=2	4	3=2	3=2	4

Here,  $\mathbf{1} = \mathbf{2} = \mathbf{3} = \mathbf{4}$

Thus the Todd-Coexter algorithm terminates after three cosets which are same

Also, in cyclic notation  $y = (\mathbf{1})$  and  $x = (\mathbf{1})$ , so order of  $x = 1$  as well as order of  $y = 1$

Clearly  $[G : H] = 1$  since there are eight index namely  $\mathbf{1}$

$$\begin{aligned}
 |G| &= [G : H] |H| \\
 &= 1 \times 1 \\
 &= 1
 \end{aligned}$$

Therefore, the order of the above defined group is  $\mathbf{1}$ .

(g)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^4 = y^3, xy = y^2x \rangle$$

Here  $xy = y^2x$  can be re-written as  $yxy^{-1}x^{-1} = 1$

Use Todd-Coexter Algorithm

Let  $H = \langle x \rangle$  be the chosen subgroup.

Let  $i$  denote the index corresponding to the co-set which  $x$  fixes.

Now draw the table as follow

The tables are as follows

$x \quad x \quad x \quad x$

1	1	1	1	1
2	3	2	3	2
3	2	3	2	3

$y \quad y \quad y$

1	2	3	1
2	3	1	2
3	1	2	3

$y \quad y \quad x \quad y^{-1} \quad x^{-1}$

1	2	3	2	1	1
2	3	1	1	3	2
3	1	2	3	2	3

Thus the Todd-Coexter algorithm terminates after three cosets

Also, in cyclic notation  $y = (123)$  and  $x = (23)$ , so order of  $x = 2$  as well as order of  $y = 3$

Clearly  $[G : H] = 3$  since there are three index namely **1, 2, 3**

$$\begin{aligned} |G| &= [G : H] |H| \\ &= 2 \times 3 \\ &= 6 \end{aligned}$$

Therefore, the order of the above defined group is **6** .



(h)

Let  $G$  be a group given by

$$G = \langle x, y \mid x^4 = y^3, xy = y^2x \rangle$$

Here  $xy = y^2x$  can be re-written as  $yxy^{-1}x^{-1} = 1$

Use Todd-Coexter Algorithm

Let  $H = \langle x \rangle$  be the chosen subgroup.

Let  $i$  denote the index corresponding to the co-set which  $x$  fixes.

Now draw the table as follow

$x \quad x \quad x \quad x \quad x \quad x \quad x \quad x$

---



---

1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3

$y \quad y \quad y$

---



---

1	2	3	1
2	3	1	2
3	1	2	3

$x \quad x \quad y \quad x^{-1} \quad y^{-1}$

---



---

1	1	1	2	2	1
2	2	2	3	3	2
3	3	3	1	1	3

Thus the Todd-Coexter algorithm terminates after three cosets

Also, in cyclic notation  $y = (123)$  and  $x = ()$ , so order of  $x = 1$  as well as order of  $y = 3$

Clearly  $[G : H] = 3$  since there are three index namely **1, 2, 3**

$$\begin{aligned} |G| &= [G : H] |H| \\ &= 1 \times 3 \\ &= 3 \end{aligned}$$

Therefore, the order of the above defined group is **3** .

4. a

Todd coexter algorithm-Let  $G$  be a finite group with generators and relation. Let  $H$  be cyclic subgroup of  $G$ . Then following rules must be followed while using Todd Coexter Algorithm

- 1) The operation of each generator is a permutation.
- 2) The relations operate trivially: they fix every coset.
- 3) The generators of  $H$  fix coset  $[H]$ .
- 4) The operation is transitive.

[Comment](#)

Step 2 of 2 ^

Let  $G$  be a group.

Now let  $H$  to be a normal subgroup of  $G$ .

In the table the following conditions shows the normality of a subgroup  $H$  of a group  $G$ .

In the final table if for every coset  $z$  and every generator  $a$  of  $H$ ,

$$za = z$$

This implies that every generator of the subgroup when acts on any coset it makes no change in the coset.

## 5. a

A group is said to be trivial if it contains only the identity element.

Mathematically, a group  $G$  is said to be trivial if and only if  $G = \{e\}$ .

Todd coexter algorithm-Let  $G$  be a finite group with generators and relation. Let  $H$  be cyclic subgroup of  $G$ . Then following rules must be followed while using Todd Coexter Algorithm

- 1) The operation of each generator is a permutation.
- 2) The relations operate trivially: they fix every coset.
- 3) The generators of  $H$  fix coset  $[H]$ .
- 4) The operation is transitive.

Case-1

Consider the group  $G = \langle x, y \mid x^4 = 1, y^3 = 1, x^2 = yxy \rangle$

Now since  $x^2 = yxy$

$$x^2 = yxy$$

$$x^4 = yxy yxy$$

$$= yxy^2xy$$

$$= 1$$

Also  $x^2 = yxy$  implies  $x^4 = xyxyx = 1$

Thus  $xyxyx = yxy^2xy = 1$

Now  $(yx)^{-1} = y^2xy$  and  $(yx) = xyx$

So,  $y^2xy = xyx$

Premultiply by  $y$  on both the sides and use the fact that  $x^3 = yxyx$

$$y^2xy = xyx$$

$$xy = yxyx$$

$$= x^2x$$

$$= x^3$$

Premultiply on both sides by  $x$

$$x^3 = xy$$

$$x^4 = x^2y$$

$$1 = x^2y$$

Now  $1 = x^2y$  implies that  $y = (x^2)^{-1}$  and  $x^2$  is the inverse of itself so

$$y = x^2$$

Since  $y^3 = 1$  implies  $x^6 = 1$  and since  $x^4 = 1$  thus  $x^2 = 1$

Hence  $1 = x^2y$  and  $x^2 = 1$  implies that  $y = 1$

Also  $x^2 = yxy$  and  $y = 1$  implies that  $x^2 = x$

Also

$$\begin{aligned} x^3 &= x^2x \\ &= x \end{aligned}$$

Thus  $x^3 = x = x^2 = 1$

Hence,  $x = 1 = y$

Therefore the group is trivial only.

Case2

Consider the group  $G = \langle x, y \mid x^4 = 1, y^3 = 1, x^2 = yxy \rangle$

Use Todd-Coexter Algorithm

Let  $H = \langle x \rangle$  be the chosen subgroup.

Let **1** denote the index corresponding to the co-set which  $x$  fixes.

Now draw the table as follow

$x \quad x \quad x \quad x$

---



---

<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>

$y \quad y \quad y$

---



---

<b>1</b>	<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>1</b>
<b>3 = 4 = 2</b>	<b>3 = 4</b>	<b>1</b>	<b>3 = 4 = 2</b>

---



---

$x \quad x \quad y \quad x \quad y$

<b>1</b>	<b>1</b>	<b>1</b>	<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>1</b>
<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>3 = 4 = 2</b>	<b>1</b>	<b>1</b>	<b>3 = 4 = 2</b>

Now use Todd coexter algorithm to deduce following relation

Since **4** under  $y$  goes to **1** but also **3** under  $y$  goes to **1**

Hence **3 = 4**

Also since **3** under  $x$  goes to **2** as well as **4 = 3**, so **3 = 2**

Since **3** under  $y$  goes to **1** as well as **3**

This implies that **3 = 2 = 1 = 4**

Thus there is only one index for the subgroup  $H = \langle x \rangle$ .

Also, in cyclic notation  $x = (1)$  and  $y = (1)$ , so order of  $x = 1$  as well as order of  $y = 1$

Clearly  $[G : H] = 1$  since only one index remains, thus

$$\begin{aligned} |G| &= [G : H] |H| \\ &= 1 \end{aligned}$$

Hence  $G$  is a group whose order is one. Thus is the trivial group.

**Therefore, the group with representation  $G = \langle x, y \mid x^4 = 1, y^3 = 1, x^2 = yxy \rangle$  is the trivial group.**

6. a

Todd coexter algorithm-Let  $G$  be a finite group with generators and relation. Let  $H$  be cyclic subgroup of  $G$ . Then following rules must be followed while using Todd Coexter Algorithm

- 1) The operation of each generator is a permutation.
- 2) The relations operate trivially: they fix every coset.
- 3) The generators of  $H$  fix coset  $[H]$ .
- 4) The operation is transitive.

Let  $G^{pqr} = \langle x, y, z \mid x^p, y^q, z^r, xyz \rangle$  denote a triangle group where  $p \leq q \leq r$

(a)

Let  $p = 2$ ,  $q = 2$  and  $r = n$

So  $G^{22n} = \langle x, y, z \mid x^2, y^2, z^n, xyz \rangle$

Now since  $xyz = 1$ , pre-multiply on both sides by  $x$

$$xyz = 1$$

$$x^2 yz = x$$

$$yz = x$$

Since  $x^2 = 1$  thus  $yz yz = 1$

Hence  $G^{22n} = \langle y, z \mid y^2, z^n, (yz)^2 \rangle$  and since representation of dihedral group of order  $n$  is given by

$$D_n = \langle p, q \mid p^2, q^n, (pq)^2 \rangle$$

Compare both the groups

$$G^{22n} = \langle y, z \mid y^2, z^n, (yz)^2 \rangle \cong D_n = \langle p, q \mid p^2, q^n, (pq)^2 \rangle$$

**Therefore, the triangle group  $G^{pqr} = \langle x, y, z \mid x^p, y^q, z^r, xyz \rangle$  with  $p = 2$ ,  $q = 2$  and  $r = n$  is isomorphic to the dihedral group  $D_n$  of order  $n$**

(b)

Let  $p = 2$ ,  $q = 3$  and  $r = 4$

So  $G^{234} = \langle x, y, z \mid x^2, y^3, z^4, xyz \rangle$

Now since  $xyz = 1$ , pre-multiply on both sides by  $x$

$xyz = 1$

$x^2yz = x$

$yz = x$

Since  $x^2 = 1$  thus  $yzyz = 1$

Thus  $G^{234} = \langle y, z \mid y^2, z^4, (yz)^2 \rangle$

Let  $H = \langle z \rangle$  be a subgroup generated by  $z$ .

Use Todd-Coexter Algorithm to obtain order of  $H = \langle z \rangle$ .

Let  $\mathbf{1}$  denote the index corresponding to the co-set which  $x$  fixes.

Now follow the given table

$y \quad y \quad y$

1	2	3	1
2	3	1	2
3	1	2	3
4	5	6	4
5	6	4	5
6	4	5	6

$z \quad z \quad z \quad z$

1	1	1	1	1
2	3	4	5	2
3	4	5	2	3
4	5	2	3	4
5	2	3	4	5
6	6	6	6	6

$y \quad z \quad y \quad z$

1	2	3	1	1
2	3	4	5	2
3	1	1	2	3
4	5	2	3	4
5	6	6	4	5
6	4	5	6	6

Since  $[G : H] = 6$ , as the number of indices in the table is 6.

And from the above table the permutation representation of  $z$  and  $y$  is given by

$$z = (2345)$$

$$y = (123)(456)$$

Hence the order of  $z$  is 4 and thus  $|H| = 4$

Since,  $|G| = [G : H]|H|$  holds

Hence,

$$\begin{aligned} |G| &= [G : H]|H| \\ &= 4 \times 6 \\ &= 24 \end{aligned}$$

Since representation of octahedral group of order 24 is given by

$$O = \langle p, q \mid p^2, q^3, (pq)^4 \rangle$$

As  $|G| = |O| = 24$ , thus the groups are isomorphic

**Therefore, the triangle group  $G^{pqr} = \langle x, y, z \mid x^p, y^q, z^r, xyz \rangle$  with  $p = 2$ ,  $q = 3$  and  $r = 4$  is isomorphic to the octahedral group  $O$ .**

(c)

Let  $p = 2$ ,  $q = 3$  and  $r = 5$

$$\text{So } G^{235} = \langle x, y, z \mid x^2, y^3, z^5, xyz \rangle$$

Now since  $xyz = 1$ , pre-multiply on both sides by  $x$

$$xyz = 1$$

$$x^2yz = x$$

$$yz = x$$

Since  $x^2 = 1$  thus  $yzyz = 1$

$$\text{Thus } G^{235} = \langle y, z \mid y^3, z^5, (yz)^2 \rangle$$

Let  $H = \langle z \rangle$  be a subgroup generated by  $z$ .

Use Todd-Coexter Algorithm to obtain order of  $H = \langle z \rangle$ .

Let  $\mathbf{1}$  denote the index corresponding to the co-set which  $x$  fixes.

Now follow the given table



y y y

1	2	3	1
2	3	1	2
3	1	2	3
4	6	8	4
5	7	9	5
6	8	4	6
7	9	5	7
8	4	6	8
9	5	7	9
10	11	12	10
11	12	10	11
12	10	11	12

---

---

1	1	1	1	1	1
2	3	4	5	6	2
3	4	5	6	2	3
4	5	6	2	3	4
5	6	2	3	4	5
6	2	3	4	5	6
7	8	9	10	11	7
8	9	10	11	7	8
9	10	11	7	8	9
10	11	7	8	9	10
11	7	8	9	10	11
12	12	12	12	12	12

$y \ z \ y \ z$

---



---

1	2	3	1	1
2	3	4	6	2
3	1	1	2	3
4	6	2	3	4
5	7	8	4	5
6	8	9	5	6
7	9	10	11	7
8	4	5	7	8
9	4	5	8	9
10	11	7	9	10
11	12	12	10	11
12	10	11	12	12

Since  $[G : H] = 12$ , as the number of indices in the table is 12.

And from the above table the permutation representation of  $z$  and  $y$  is given by

$$z = (23456)(7891011)$$

$$y = (123)(468)(579)(101112)$$

Hence order of  $z$  is 5 and thus  $|H| = 5$

Since,  $|G| = [G : H]|H|$  holds

Hence,

$$\begin{aligned} |G| &= [G : H]|H| \\ &= 12 \times 5 \\ &= 60 \end{aligned}$$

Since representation of icosahedral group of order 60 is given by

$$I = \langle p, q \mid p^2, q^3, (pq)^5 \rangle$$

As  $|G| = |I| = 60$ , thus the groups are isomorphic

**Therefore, the triangle group  $G^{pqr} = \langle x, y, z \mid x^p, y^q, z^r, xyz \rangle$  with  $p = 2$ ,  $q = 3$  and  $r = 5$  is isomorphic to the icosahedral group  $I$ .**

7. a

!!!

8. a

!!!

# Miscellaneous Problem

1. a

Consider the groups that are generated by two elements  $x$  and  $y$  of order 2.

Since,

The elements  $x$  and  $y$  have order 2,

If the generated group is abelian then,

$$\begin{aligned}(xy)^2 &= xyxy \\ &= xxyy \\ &= x^2y^2 \\ &= 1\end{aligned}$$

So, the generated group will be Klein four group.

Now,

Consider the generated group  $G$  is not abelian,

Then,

Assume,

$$x = b, y = ab$$

As,

$$x^2 = 1, y^2 = 1$$

So,

$$b^2 = 1, (ab)^2 = abab = 1$$

Then,

$$a = (ab)b^{-1}$$

Therefore,

By closure property,

$$a \in G$$

Let,

The order of  $a$  is  $n$ ,

Then

$$\begin{aligned}a^n &= (ab)b^{-1}(ab)b^{-1}(ab)b^{-1} \dots (ab)b^{-1} \\ &= a(bb^{-1})a(bb^{-1})a(bb^{-1}) \dots a(bb^{-1}) \\ &= aaa \dots a \\ &= 1\end{aligned}$$

Now,

The dihedral group  $D_n$  is isomorphic to the below group:

$$\langle a, b : a^n, b^2, abab \rangle$$

Therefore, the group that is generated by two elements  $x$  and  $y$  of order 2 is either Klein four abelian group or the non-abelian dihedral group  $D_n$ .

2. a

Let  $D_n$  be a dihedral group of order  $2n$ . Then  $D_n$  is generated by two elements say  $x$  and  $y$  that satisfy the following conditions

$$x^n = 1, y^2 = 1 \text{ and } yx = x^{-1}y$$

$$\text{Thus } D_n = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$$

Double Coset of a group with respect two subgroups  $H, K$  is denoted by  $HgK$  where  $g$  is an arbitrary element of the group and is defined as follows

$$HgK = \{h g k : h \in H, k \in K\}$$

Let  $D_n$  be a dihedral group. Let  $H = \{1, y\}$  be a subgroup of  $D_n$ .

The possible double coset  $HgH$  are given as follows

$$\{g, yg, gy, ygy\} \text{ where } g \in D_n \text{ denotes an arbitrary element}$$

#### Case 1

For  $HgH = g$

Clearly the whole double coset is equal to the space  $D_n$ .

#### Case 2

For  $HgH = yg$

Let  $g \in D_n$  be an arbitrary element then  $g \in \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

If  $g = x^i, 0 \leq i \leq n-1$  then

$$\begin{aligned} yg &= yx^i \\ &= x^{-i}y \end{aligned}$$

If  $g = y$  then

$$\begin{aligned} yg &= y^2 \\ &= 1 \end{aligned}$$

If  $g = x^i y, 0 \leq i \leq n-1$

$$\begin{aligned} yg &= yx^i y \\ &= yx(x)^{i-1} y \\ &= x^{-1}y(x)^{i-1} y \end{aligned}$$

Continuous simplification as above implies that  $yg = x^i$

$$\text{Thus } HgH = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\} = D_n$$

#### Case 3

For  $HgH = gy$

This case is same as the second one.

$$\text{Hence, } HgH = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\} = D_n$$

#### Case 4

For  $HgH = ygy$

Let  $g \in D_n$  be an arbitrary element then  $g \in \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

If  $g = x^i, 0 \leq i \leq n-1$  then

$$\begin{aligned} ygy &= yx^i y \\ &= yx(x)^{i-1} y \\ &= x^{-1}y(x)^{i-1} y \end{aligned}$$

Continuous simplification as above implies that  $ygy = x^i$

If  $g = y$  then  $ygy = y^3 = 1$

If  $g = x^i y, 0 \leq i \leq n-1$  then,

$$\begin{aligned} y(x^i y)y &= yx^i y^2 \\ &= yx^i \\ &= x^{-i}y \end{aligned}$$

Thus in this case also

$$HgH = D_n$$

Now use the formula  $|HgK| = |H| \cdot [K : K \cap g^{-1}Hg]$

Since here  $H = K$ ,

$$\text{So } |HgH| = |H| \cdot [H : H \cap g^{-1}Hg]$$

Now

$$[H : H \cap g^{-1}Hg] = \frac{|H|}{|H \cap g^{-1}Hg|}$$

Since  $|H| = 2$  and  $|H \cap g^{-1}Hg| = \{1, 2\}$

Hence  $[H : H \cap g^{-1}Hg] = \{1, 2\}$

Thus  $|HgH| = \{2, 4\}$

Therefore the double coset of the subgroup  $H = \{1, y\}$  of the dihedral groups  $D_n$  are the group itself and the number of elements in double coset of the subgroup  $H = \{1, y\}$  is either 2 or 4 .

### 3. a

Let  $D_n$  be a dihedral group of order  $2n$ . Then  $D_n$  is generated by two elements say  $x$  and  $y$  that satisfy the following conditions

$$x^n = 1, y^2 = 1 \text{ and } yx = x^{-1}y$$

Thus  $D_n = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

Double Coset of a group with respect two subgroups  $H, K$  is denoted by  $HgK$  where  $g$  is an arbitrary element of the group and is defined as follows

$$HgK = \{h g k : h \in H, k \in K\}$$



Let  $D_n$  be a dihedral group. Let  $H = \{1, y\}$  be a subgroup of  $D_n$ .

The possible double coset  $HgH$  are given as follows

$\{g, yg, gy, ygy\}$  where  $g \in D_n$  denotes an arbitrary element

#### Case 1

For  $HgH = g$

Clearly the whole double coset is equal to the space  $D_n$ .

#### Case 2

For  $HgH = yg$

Let  $g \in D_n$  be an arbitrary element then  $g \in \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

If  $g = x^i$ ,  $0 \leq i \leq n-1$  then

$$\begin{aligned} yg &= yx^i \\ &= x^{-i}y \end{aligned}$$

If  $g = y$  then

$$\begin{aligned} yg &= y^2 \\ &= 1 \end{aligned}$$

If  $g = x^i y$ ,  $0 \leq i \leq n-1$

$$\begin{aligned} yg &= yx^i y \\ &= yx(x)^{i-1} y \\ &= x^{-1}y(x)^{i-1} y \end{aligned}$$

Continuous simplification as above implies that  $yg = x^i$

Thus  $HgH = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\} = D_n$

#### Case 3

For  $HgH = gy$

This case is same as the second one.

Hence,  $HgH = \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\} = D_n$

#### Case 4

For  $HgH = ygy$

Let  $g \in D_n$  be an arbitrary element then  $g \in \{1, x, x^2, \dots, x^{n-1}, y, xy, x^2y, \dots, x^{n-1}y\}$

If  $g = x^i$ ,  $0 \leq i \leq n-1$  then

$$\begin{aligned} ygy &= yx^i y \\ &= yx(x)^{i-1} y \\ &= x^{-1}y(x)^{i-1} y \end{aligned}$$

Continuous simplification as above implies that  $ygy = x^i$

If  $g = y$  then  $ygy = y^3 = 1$

If  $g = x^i y$ ,  $0 \leq i \leq n-1$  then,

$$\begin{aligned} y(x^i y)y &= yx^i y^2 \\ &= yx^i \\ &= x^{-i}y \end{aligned}$$

Thus in this case also

$$HgH = D_n$$

Now use the formula  $|HgK| = |H| \cdot [K : K \cap g^{-1}Hg]$

Since here  $H = K$ ,

So  $|HgH| = |H| \cdot [H : H \cap g^{-1}Hg]$

Now

$$[H : H \cap g^{-1}Hg] = \frac{|H|}{|H \cap g^{-1}Hg|}$$

Since  $|H| = 2$  and  $|H \cap g^{-1}Hg| = \{1, 2\}$

Hence  $[H : H \cap g^{-1}Hg] = \{1, 2\}$

Thus  $|HgH| = \{2, 4\}$

Therefore the double coset of the subgroup  $H = \{1, y\}$  of the dihedral groups  $D_n$  are the group itself and the number of elements in double coset of the subgroup  $H = \{1, y\}$  is either 2 or 4.

4. a

Consider  $H$  and  $K$  be subgroups of a group  $G$ , with  $H \subset K$ .

And,

Suppose that  $H$  is normal in  $K$ , and that  $K$  is normal in  $G$ .

[Comment](#)

Step 2 of 2 ^

Since,  $H$  is normal in  $K$ .

Then,

For some  $k \in K$  and  $h \in H$ ,

$$khk^{-1} \in K$$

Also,  $K$  is normal in  $G$ .

So,

For some  $g \in G$  and  $khk^{-1} \in K$ ,

$$g(khk^{-1})g^{-1} \in G$$

Thus,

$$(gk)h(gk)^{-1} \in G$$

As,  $g \in G$  and  $k \in K \subset G$ ,

Therefore,

By closure property,  $gk \in G$

Hence,  $H$  is normal in  $G$ .

5. a

First Isomorphism theorem-Let  $G$  and  $H$  be groups and  $\phi: G \rightarrow H$  be a homomorphism, then the image of the homomorphism is isomorphic to the quotient group  $G/\ker(\phi)$ .

Mathematically,

$$\phi(G) \cong G/\ker(\phi)$$

(a)

Let  $\pi : G \rightarrow G/N$  be the canonical homomorphism. Let  $H, N$  be subgroups of  $G$

Restrict this map to  $H$ , thus

$$\pi : H \rightarrow H/N \text{ defined by } \pi(h) = hN$$

$$\text{Now } \ker(\pi) = \{h \in H : \pi(h) = N\}$$

$$\begin{aligned}\ker(\pi) &= \{h \in H : \pi(h) = N\} \\ &= \{h \in H : hN = N\} \\ &= \{h \in H : h \in N\}\end{aligned}$$

$$\text{Hence } \ker(\pi) = H \cap N$$

Now restrict the map  $\pi : G \rightarrow G/N$  to  $HN$

$$\pi : HN \rightarrow HN/N \text{ defined by } \pi(p) = pN, \text{ where } p \in HN \text{ is arbitrary.}$$

$$\text{Now } \ker(\pi) = \{p \in HN : \pi(p) = N\}$$

$$\begin{aligned}\ker(\pi) &= \{p \in HN : \pi(p) = N\} \\ &= \{p = hn, h \in H, n \in N : pN = N\} \\ &= \{p \in HN : hN = N\}\end{aligned}$$

This implies that  $h \in N$ , thus  $p \in N$

Hence,  $\ker(\pi) = N$

**Therefore, for  $\pi : G \rightarrow G/N$  homomorphism the kernel of restriction of this map to  $H$  is  $H \cap N$  and the kernel of restriction of this map to  $HN$  is  $N$ .**

(b)

Since  $\pi : H \rightarrow H/N$  is surjective. By first isomorphism theorem,

$$\begin{aligned}H/N &\cong H/\ker(\pi) \\ H/N &\cong H/H \cap N\end{aligned}$$

Again the restriction of the map  $\pi : G \rightarrow G/N$  to  $HN$  is also surjective, By first isomorphism theorem,

$$\begin{aligned}HN/N &\cong HN/\ker(\pi) \\ HN/N &\cong H/N\end{aligned}$$

Since  $H/N \cong H/H \cap N$  and  $HN/N \cong H/N$

Thus  $H/H \cap N \cong HN/N$

**Therefore, the third isomorphism theorem holds in the above case that is**

$$H/H \cap N \cong HN/N.$$

6. a

Let  $G$  be a group. A subgroup  $H$  of  $G$  is said to be normal if following condition holds

For every  $x \in G$ ,  $xHx^{-1} = H$

First Isomorphism Theorem- Let  $G$  and  $H$  be a group such that there exists a homomorphism between them then the kernel of the homomorphism is a normal subgroup of  $G$ .

(a)

Let  $G$  be a group and  $H, N$  be normal subgroups of  $G$  such that  $N \subset H$

Define a map

$$\phi: \frac{G}{N} \rightarrow \frac{G}{H} \text{ by } \phi(gN) = gH \text{ for an arbitrary element } g \in G$$

Let  $g', g \in G$  be arbitrary elements

Let  $g'N = gN$  this implies that  $g' = gn$ , for some  $n \in N$

Since  $N \subset H$ , thus  $n \in N \subset H$

Thus,  $g'H = gH$

This implies that  $\phi(g'N) = \phi(gN)$ . Thus the map is well-defined.

Also  $g'NgN = g'gN$

Now apply the map on the element  $g'NgN$  so

$$\begin{aligned}\phi(g'NgN) &= \phi(g'gN) \\ &= g'gH \\ &= g'HgH\end{aligned}$$

Hence above defined map is a homomorphism

By First Isomorphism Theorem,  $\ker(\phi)$  is a normal subgroup of  $G/N$ .

Now  $\ker(\phi)$  is given by  $\ker(\phi) = \{gN : \phi(gN) = N\}$ .

Clearly the whole space  $G/N = \ker(\phi)$

Also the map is surjective.

Thus  $G/N$  is a normal subgroup of  $G/H$

**Therefore for a group  $G$  and  $H, N$  be normal subgroups of  $G$  such that  $N \subset H$ ,**

**$\overline{H} = H/N$  is a normal subgroup of  $\overline{G} = G/N$ .**

(b)

Consider the homomorphism  $G \rightarrow \overline{G} \rightarrow \overline{G}/\overline{H}$  defined by

$$g \rightarrow gN \rightarrow (gN)\overline{H}, \text{ where } g \in G \text{ is an arbitrary element.}$$

Consider the map  $\overline{G} \rightarrow \overline{G}/\overline{H}$  defined by  $gN \rightarrow (gN)\overline{H}$ .

Let  $(g_1N)\overline{H} = (g_2N)\overline{H}$ , then for some  $h \in H$

$$\begin{aligned}(g_1N)\overline{H} &= (g_2N)\overline{H} \\ g_1N &= (g_2N)h \\ g_1N &= g_2N\end{aligned}$$

Hence the map is injective.

Also for every element  $(gN)\overline{H} \in \overline{G}/\overline{H}$  there exists an element  $gN \in \overline{G}$  such that

$$gN \rightarrow (gN)\overline{H}.$$

Thus the map is surjective also. Thus  $\overline{G} \rightarrow \overline{G}/\overline{H}$  is a bijective homomorphism.

Similarly  $G \rightarrow \overline{G}$  is also a bijective homomorphism.

Thus  $G \rightarrow \overline{G} \rightarrow \overline{G}/\overline{H}$  is a bijective homomorphism.

This implies that  $G \cong \overline{G}/\overline{H}$ .

Now define a map  $G \rightarrow G/H$  by  $g \rightarrow gH$

Clearly this map is surjective as well as a homomorphism

Let  $g_1 \neq g_2$ , where  $g_1, g_2 \in G$

Since  $g_1H \neq g_2H$  holds, hence the above defined map is injective also.

This map is also a bijective homomorphism.

Hence following holds,

$$\overline{G}/\overline{H} \cong G \cong G/H$$

**Therefore, the third isomorphism theorem holds here that is  $\overline{G}/\overline{H} \cong G/H$ .**

7. a

Consider  $p_1, p_2$  be permutations of the set  $S = \{1, 2, \dots, n\}$  and let  $U_i$  be the subset of  $S$  of indices that are not fixed by  $p_i$ .

---

[Comment](#)

---

Step 2 of 3 ^

(a)

Since,

$$U_1 \cap U_2 = \emptyset$$

Therefore, the elements that are not fixed by  $p_1, p_2$  are different.

So,

The commutator  $p_1 p_2 p_1^{-1} p_2^{-1}$  will follow commutative law as multiplication of disjoint permutation.

Therefore, if  $U_1 \cap U_2 = \emptyset$  then the commutator  $p_1 p_2 p_1^{-1} p_2^{-1}$  is the identity.

(b)

Consider,

The set  $U_1 \cap U_2$  contains exactly one element  $a$ .

So,

The commutator  $p_1 p_2 p_1^{-1} p_2^{-1}$  will be a three cycle containing  $a$ ,

Because,

Both of the permutations  $p_1$  and  $p_2$  are not fixing it while other elements are fixed by  $p_1$  and  $p_2$ .

Therefore, if  $U_1 \cap U_2$  contains exactly one element then the commutator  $p_1 p_2 p_1^{-1} p_2^{-1}$  will be a three cycle.

8. a

Consider  $H$  be a subgroup of an infinite group  $G$ .

Assume,

The number of left cosets is  $n_1$  and the number of right cosets is  $n_2$ .

Since,

The left cosets of a subgroup  $H$  of an infinite group  $G$  will partition the group.

And,

A partition is a subdivision of  $G$  into non-overlapping, non-empty subgroups.

So,

$$G = \{\cup_a H : a \in G\}$$

Thus,

$$[G : H] = n_1$$

Also,

The right cosets of a subgroup  $H$  of an infinite group  $G$  will partition the group.

Therefore,

$$G = \{\cup_b H_b : b \in G\}$$

Then,

$$[G : H] = n_2$$

As a result,

$$n_1 = n_2$$

Hence, **the number of left cosets is equal to the number of right cosets also when  $G$  is an infinite group.**

## 9. a

Consider  $x$  be an element, not the identity  $e$ , of a group  $K$  of odd order  $2n+1$  where  $n \in \mathbb{N}$ .

Then,

For any  $y \in K$ ,  $y^{2n+1} = e$ .

Assume,

The elements  $x$  and  $x^{-1}$  are conjugate.

Then,

For some  $y \in K$ ,

$$yxy^{-1} = x^{-1}$$

After taking inverse on both sides,

$$(y^{-1})^{-1} x^{-1} y^{-1} = (x^{-1})^{-1}$$

$$yx^{-1}y^{-1} = x$$

Now,

With the help of above two results,

$$\begin{aligned} y^2xy^{-2} &= y(yxy^{-1})y^{-1} \\ &= yx^{-1}y^{-1} \\ &= x \end{aligned}$$

Thus,

$$\begin{aligned} y^3xy^{-3} &= y(y^2xy^{-2})y^{-1} \\ &= yxy^{-1} \\ &= x^{-1} \end{aligned}$$

As,

$$\begin{aligned} y^4xy^{-4} &= y(y^3xy^{-3})y^{-1} \\ &= yx^{-1}y^{-1} \\ &= x \end{aligned}$$



And,

$$\begin{aligned} y^5 xy^{-5} &= y(y^4 xy^{-4})y^{-1} \\ &= yxy^{-1} \\ &= x^{-1} \end{aligned}$$

Also,

$$\begin{aligned} y^6 xy^{-6} &= y(y^5 xy^{-5})y^{-1} \\ &= yx^{-1}y^{-1} \\ &= x \end{aligned}$$

Further,

$$\begin{aligned} y^7 xy^{-7} &= y(y^6 xy^{-6})y^{-1} \\ &= yxy^{-1} \\ &= x^{-1} \end{aligned}$$

Similarly,

For odd natural number  $2n-1$ ,

$$y^{2n-1} xy^{-(2n-1)} = x^{-1}$$

So,

For next even natural number  $2n$ ,

$$\begin{aligned} y^{2n} xy^{-2n} &= y(y^{2n-1} xy^{-(2n-1)})y^{-1} \\ &= yx^{-1}y^{-1} \\ &= x \end{aligned}$$

Therefore,

For odd natural number  $|K| = 2n+1$ ,

$$\begin{aligned} y^{2n+1} xy^{-(2n+1)} &= y(y^{2n} xy^{-2n})y^{-1} \\ &= yxy^{-1} \\ &= x^{-1} \end{aligned}$$

Consequently,

$$\begin{aligned} exe^{-1} &= x^{-1} \\ x &= x^{-1} \\ x^2 &= e \end{aligned}$$

As a result,

The element  $x$  has even order 2,

But  $|x|$  should be a divisor of  $|K| = 2n+1$ ,

So,

Our assumption, the elements  $x$  and  $x^{-1}$  are conjugate, is wrong.

Hence, **the non-identity elements  $x$  and  $x^{-1}$  are not conjugate in a group of odd order.**

10. a

Consider  $G$  be a finite group that operates transitively on a set  $S$  of order  $\geq 2$ .

[Comment](#)

Step 2 of 2 ^

Since,

The group  $G$  is a finite group that operates transitively on a set  $S$  of order  $\geq 2$ .

Then,

For some  $g \in G; p_1, p_2 \in S$ ,

$$gp_1 = p_2$$

Here,

$$L.C.M.(O(g), O(p_1)) = O(p_2)$$

So,

The order of  $g$  should be a divisor of order of  $p_2 \in S$ .

If order of  $g$  is  $n$  then  $|S| - n$  fix by  $g$ .

Therefore, **if  $G$  operates transitively on a set  $S$  of order  $\geq 2$  then  $G$  should contain an element  $g$  that does not fix any element of  $S$ .**

11. a

Consider the conjugacy classes of elements of order 2 in  $GL_2(\mathbb{Z})$ .

[Comment](#)

Step 2 of 2 ^

Since,

The elements of order 2 in  $GL_2(\mathbb{Z})$  are  $\begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$  and  $\begin{bmatrix} -1 & \\ & -1 \end{bmatrix}$ .

And,

The conjugacy class of  $\begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$  contains  $\begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$  only.

Also,

The conjugacy class of  $\begin{bmatrix} -1 & \\ & -1 \end{bmatrix}$  contains  $\begin{bmatrix} -1 & \\ & -1 \end{bmatrix}$  only.

Hence, **the conjugacy classes of elements of order 2 in  $GL_2(\mathbb{Z})$  are  $\begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$  and  $\begin{bmatrix} -1 & \\ & -1 \end{bmatrix}$ .**

12. a

$SL_2(\mathbb{F}_5)$  denotes the special linear group over the finite field  $\mathbb{F}_5$ , that it is the group of matrices with determinant 1 and entries from the field  $\mathbb{F}_5$ .

Centralizer of an element of a group is the collection of all those elements of that group which commutes with that element.

(a)

Let  $A$  be a matrix in  $SL_2(\mathbb{F}_5)$  given by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & a \end{pmatrix}$$

For  $a = 0$

Centralizer of  $A = \{B \in SL_2(\mathbb{F}_5) \mid AB = BA\}$

Let

$$B = \begin{pmatrix} p & r \\ q & s \end{pmatrix} \text{ where } p, q, r, s \in \mathbb{F}_5$$

Now consider the product of these two matrices such that they commute

$$\begin{aligned} AB &= BA \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p & r \\ q & s \end{pmatrix} &= \begin{pmatrix} p & r \\ q & s \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \begin{pmatrix} -q & -s \\ p & r \end{pmatrix} &= \begin{pmatrix} r & -p \\ s & -q \end{pmatrix} \end{aligned}$$

Hence  $q = -r$  and  $s = p$ .

So  $B$  is of the form and use the modulo 5 arithmetic

$$B = \begin{pmatrix} p & -q \\ q & p \end{pmatrix} = \begin{pmatrix} p & 5-q \\ q & p \end{pmatrix}$$

Also the  $|B| = 1$  implies that  $p^2 + q^2 = 1$  under modulo 5 arithmetic

Thus the possible centralizers of  $A$  when  $a = 0$  is the set of all  $2 \times 2$  matrices  $B$  of the form

$$\begin{pmatrix} p & 5-q \\ q & p \end{pmatrix} \text{ such that } p^2 + q^2 = 1.$$

For  $a = 1$

Centralizer of  $A = \{B \in SL_2(\mathbb{F}_5) \mid AB = BA\}$

Let

$$B = \begin{pmatrix} p & r \\ q & s \end{pmatrix} \text{ where } p, q, r, s \in \mathbb{F}_5$$

Now consider the product of these two matrices such that they commute

$$\begin{aligned} AB &= BA \\ \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} p & r \\ q & s \end{pmatrix} &= \begin{pmatrix} p & r \\ q & s \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \\ \begin{pmatrix} -q & -s \\ p+q & r+s \end{pmatrix} &= \begin{pmatrix} r & -p+r \\ s & -q+s \end{pmatrix} \end{aligned}$$

Hence  $q = -r$  and  $p+q = s$

So  $B$  is of the form and use the modulo 5 arithmetic

$$B = \begin{pmatrix} p & -q \\ q & p+q \end{pmatrix} = \begin{pmatrix} p & 5-q \\ q & p+q \end{pmatrix}$$

Also the  $|B| = 1$  implies that  $p^2 + q^2 + pq = 1$  under modulo 5 arithmetic

Thus the possible centralizers of  $A$  when  $a = 1$  is the set of all  $2 \times 2$  matrices  $B$  of the form

$$\begin{pmatrix} p & 5-q \\ q & p+q \end{pmatrix} \text{ such that } p^2 + q^2 + pq = 1 \text{ under modulo 5 arithmetic}$$

For  $a = 2$

Centralizer of  $A = \{B \in SL_2(\mathbb{F}_5) \mid AB = BA\}$

Let

$$B = \begin{pmatrix} p & r \\ q & s \end{pmatrix} \text{ where } p, q, r, s \in \mathbb{F}_5$$

Now consider the product of these two matrices such that they commute

$$\begin{aligned} AB &= BA \\ \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} p & r \\ q & s \end{pmatrix} &= \begin{pmatrix} p & r \\ q & s \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \\ \begin{pmatrix} -q & -s \\ p+2q & r+2s \end{pmatrix} &= \begin{pmatrix} r & -p+2r \\ s & -q+2s \end{pmatrix} \end{aligned}$$

Hence  $q = -r$  and  $p+2q = s$

So  $B$  is of the form and use the modulo 5 arithmetic

$$B = \begin{pmatrix} p & -q \\ q & p+2q \end{pmatrix} = \begin{pmatrix} p & 5-q \\ q & p+2q \end{pmatrix}$$

Also the  $|B| = 1$  implies that  $p^2 + q^2 + 2pq = 1$  under modulo 5 arithmetic

Thus the possible centralizers of  $A$  when  $a = 0$  is the set of all  $2 \times 2$  matrices  $B$  of the form

$$\begin{pmatrix} p & 5-q \\ q & p+2q \end{pmatrix} \text{ such that } (p+q)^2 = 1 \text{ under modulo 5 arithmetic}$$

For  $a = 3$

Centralizer of  $A = \{B \in SL_2(\mathbb{F}_5) \mid AB = BA\}$

Let

$$B = \begin{pmatrix} p & r \\ q & s \end{pmatrix} \text{ where } p, q, r, s \in \mathbb{F}_5$$

Now consider the product of these two matrices such that they commute

$$\begin{aligned} AB &= BA \\ \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} p & r \\ q & s \end{pmatrix} &= \begin{pmatrix} p & r \\ q & s \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \\ \begin{pmatrix} -q & -s \\ p+3q & r+3s \end{pmatrix} &= \begin{pmatrix} r & -p+3r \\ s & -q+3s \end{pmatrix} \end{aligned}$$

Hence  $q = -r$  and  $p+3q = s$

So  $B$  is of the form and use the modulo 5 arithmetic

$$B = \begin{pmatrix} p & -q \\ q & p+3q \end{pmatrix} = \begin{pmatrix} p & 5-q \\ q & p+3q \end{pmatrix}$$

Also the  $|B| = 1$  implies that  $p^2 + q^2 + 3pq = 1$  under modulo 5 arithmetic

Thus the possible centralizers of  $A$  when  $a = 0$  is the set of all  $2 \times 2$  matrices  $B$  of the form

$$\begin{pmatrix} p & 5-q \\ q & p+3q \end{pmatrix} \text{ such that } p^2 + q^2 + 3pq = 1 \text{ under modulo 5 arithmetic}$$

For  $a = 4$

Centralizer of  $A = \{B \in SL_2(\mathbb{F}_5) \mid AB = BA\}$

Let

$$B = \begin{pmatrix} p & r \\ q & s \end{pmatrix} \text{ where } p, q, r, s \in \mathbb{F}_5$$

Now consider the product of these two matrices such that they commute

$$\begin{aligned} AB &= BA \\ \begin{pmatrix} 0 & -1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} p & r \\ q & s \end{pmatrix} &= \begin{pmatrix} p & r \\ q & s \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 4 \end{pmatrix} \\ \begin{pmatrix} -q & -s \\ p+4q & r+4s \end{pmatrix} &= \begin{pmatrix} r & -p+4r \\ s & -q+4s \end{pmatrix} \end{aligned}$$

Hence  $q = -r$  and  $p + 4q = s$

So  $B$  is of the form and use the modulo 5 arithmetic

$$B = \begin{pmatrix} p & -q \\ q & p+4q \end{pmatrix} = \begin{pmatrix} p & 5-q \\ q & p+4q \end{pmatrix}$$

Also the  $|B| = 1$  implies that  $p^2 + q^2 + 4pq = 1$  under modulo 5 arithmetic

Thus the possible centralizers of  $A$  when  $a = 0$  is the set of all  $2 \times 2$  matrices  $B$  of the form

$$\begin{pmatrix} p & 5-q \\ q & p+4q \end{pmatrix} \text{ such that } p^2 + q^2 + 4pq = 1 \text{ under modulo 5 arithmetic}$$

(b)

Consider the group  $SL_2(\mathbb{F}_5)$ .

Now in order to obtain all the conjugacy classes start with evaluating the characteristic polynomial corresponding to the different eigenvalues.

Case1-Scalar Conjugacy Classes

$x^2 - 2x + 1$  or  $x^2 + 2x + 1$  is the characteristic polynomial for conjugacy classes

Here the minimal polynomial is given by  $x - 1$  or  $x + 1$

The eigenvalues are given by the solution of the equation  $(x-1)^2 = 0$  and  $(x+1)^2 = 0$

Hence the possible set of eigenvalues is given by  $\{1, -1\}$ .

Thus the number of such conjugacy classes is 2.

Thus the total number of elements is 2.

In the tabular form

Nature of conjugacy classes	Eigenvalues	Size of conjugacy class	No of conjugacy classes	No of elements
Scalar	$\{1, 1\}, \{-1, -1\}$	1	2	2
Matrix with distinct diagonal entries	$\{2, 3\}$	30	1	30
Diagonalization over $\mathbb{F}_{25}$	$\{2 + \sqrt{3}, 2 - \sqrt{3}\}$	20	2	40
Matrix which are not diagonal	$\{1, 1\}, \{-1, -1\}$	12	4	48

Hence the class equation is given by

$$120 = 1 + 1 + 12 + 12 + 12 + 12 + 20 + 20 + 30$$

Therefore the class equation of  $SL_2(\mathbb{F}_5)$  is given by

$$120 = 1 + 1 + 12 + 12 + 12 + 12 + 20 + 20 + 30.$$

(c)

Let  $\mathbb{F}_p$  be a finite field

Consider the equation  $x^2 + axy + y^2 = 1$ ,

Now let  $y = \lambda x + 1$ ,

Substitute the value  $y = \lambda x + 1$  in  $x^2 + axy + y^2 = 1$

Hence,

$$x^2 + ax(\lambda x + 1) + (\lambda x + 1)^2 = 1$$

This implies that  $x = -(2\lambda + a)/(\lambda^2 + a\lambda + 1)$ .

Thus the value of  $y$  is given by

$$y = \left( -\lambda(2\lambda + a)/(\lambda^2 + a\lambda + 1) \right) + 1$$

$$y = \frac{1 - \lambda^2}{(\lambda^2 + a\lambda + 1)}$$

Thus the solution is the point  $P$  given by

$$P = \left( \frac{1 - \lambda^2}{(\lambda^2 + a\lambda + 1)}, \frac{-(2\lambda + a)}{(\lambda^2 + a\lambda + 1)} \right)$$

First, note that if two points are given by the parametrization  $y = \lambda x + 1$  with  $x \neq 0$  and different values of  $\lambda$ , then they are distinct points.

Now, consider the case of  $\lambda^2 + a\lambda + 1 = 0$

This has at most 2 roots in  $\mathbb{F}_p$ .

In addition, the point  $(0, -1)$  not described by the parametrization.

Consequently, there are between  $(p-2)+1 = p-1$  and  $p+1$  points on the curve

$x^2 + axy + y^2 = 1$ , corresponding to  $\lambda^2 + a\lambda + 1$  having between 2 and 0 roots.

Thus there are either  $p-1, p+1$  or  $p$  solutions in  $\mathbb{F}_p$ .

Therefore, for a finite field  $\mathbb{F}_p$  there exists  $p-1, p+1$  or  $p$  solutions for the equation

$x^2 + axy + y^2 = 1$  in  $\mathbb{F}_p$



(d)

For the group  $SL_2(\mathbb{F}_q)$  use the same procedure to create table as done in above part

So the table has the form

Nature of conjugacy classes	Size of conjugacy class	No of conjugacy classes	No of elements
Scalar	1	2	2
Matrix with distinct diagonal entries	$q(q+1)$	$\frac{(q-3)}{2}$	$\frac{(q-3)q(q+1)}{2}$
Diagonalization over $\mathbb{F}_{25}$	$q(q-1)$	$\frac{(q-1)}{2}$	$\frac{(q-1)^2 q}{2}$
Matrix which are not diagonal	$\frac{(q^2-1)}{2}$	4	$2(q^2-1)$

Hence the class equation is given by

$$|SL_2(\mathbb{F}_q)| = 1 + 1 + [q(q+1)]_{\frac{(q-3)}{2}} + [q(q-1)]_{\frac{(q-1)}{2}} + \frac{(q^2-1)}{2} + \frac{(q^2-1)}{2} + \frac{(q^2-1)}{2} + \frac{(q^2-1)}{2}$$

Where the notation  $[r]_s$  means the term  $r$  is added  $s$  times.

Therefore the class equation of  $SL_2(\mathbb{F}_q)$  is given by

$$|SL_2(\mathbb{F}_q)| = 1 + 1 + [q(q+1)]_{\frac{(q-3)}{2}} + [q(q-1)]_{\frac{(q-1)}{2}} + \frac{(q^2-1)}{2} + \frac{(q^2-1)}{2} + \frac{(q^2-1)}{2} + \frac{(q^2-1)}{2}.$$

## Chapter 8

### Section 1

1. a

By **Exercise 5.1** of Chapter 3, we have that any  $n \times n$  real matrix can be written as a sum of a symmetric matrix (a matrix whose transpose is equal to itself) and a skew-symmetric matrix (a matrix whose transpose is equal to minus itself). A quick proof goes as follows: let  $Q$  be a real  $n \times n$  matrix, then  $Q$  can also be written as the sum

$$\frac{1}{2}(Q + Q^t) + \frac{1}{2}(Q - Q^t),$$

where from the properties  $(X + Y)^t = X^t + Y^t$  and  $(kX)^t = kX^t$ ,  $k \in \mathbb{R}$ , we immediately see that  $\frac{1}{2}(Q + Q^t)$  is symmetric and  $\frac{1}{2}(Q - Q^t)$  is skew-symmetric.

Now, as it was proved that if  $\langle \cdot, \cdot \rangle$  is a bilinear form on a vector space  $V$  with a fixed basis  $B$ , then there is a matrix  $A$  such that if  $v$  and  $w$  are vectors in  $V$  and  $X$  and  $Y$  are their coordinate vectors, respectively, then

$$\langle v, w \rangle = X^t A Y.$$

As per the first paragraph, there is a symmetric matrix  $S$  and a skew-symmetric matrix  $W$  such that  $A = S + W$ , so that we have

$$\langle v, w \rangle = X^t A Y = X^t (S + W) Y = X^t S Y + X^t W Y,$$

where we use the proposition which says that the form  $X^t A Y$  is (skew-)symmetric if and only if  $A$  is (skew-)symmetric to finish our proof.

#### Result

3 of 3

This is a consequence of the fact that any  $n \times n$  real matrix can be written as a sum of a symmetric matrix and a skew-symmetric matrix. Click for a complete proof.

### Section 2

1. a

Let  $A$  be a positive definite, symmetric real matrix. Suppose that its maximal entry is not on diagonal, say it is on the position  $(p, q)$  ( $p$ th row and  $q$ th column, where we denote that element  $a_{pq}$ ) with  $p \neq q$ ; since it is symmetric we also have that  $a_{pq} = a_{qp}$ . As it is positive definite then for any nonzero coordinate vector  $X$  we have

$$X^t A X > 0,$$

or, equivalently via multiplying everything out,

$$\sum_{1 \leq i, j \leq n} a_{ij} x_i x_j > 0 \quad (1)$$

for any choice  $x_1, \dots, x_n$  of elements of  $\mathbb{R}$ . Now obviously  $a_{pq} > 0$ , since if not then  $a_{ij} \leq 0$  for all  $i, j$  and (1) is violated for  $x_1 = x_2 = \dots = x_n = 1$ . Now let  $x_1, \dots, x_n$  be such that  $x_p = 1$ ,  $x_q = -1$ , and  $x_k = 0$  for all other  $k$ . Then we compute (1) as

$$\begin{aligned} \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j &= a_{pq} x_p x_q + a_{qp} x_q x_p + a_{pp} x_p^2 + a_{qq} x_q^2 \\ &= a_{pp} + a_{qq} - 2a_{pq} \end{aligned}$$

But now since  $a_{pq}$  is maximal then  $a_{pq} \geq a_{pp}$  and  $a_{pq} \geq a_{qq}$ , so that  $2a_{pq} \geq a_{pp} + a_{qq}$  and consequently

$$a_{pp} + a_{qq} - 2a_{pq} \leq 0$$

which is a contradiction with positive definiteness of  $A$ , thus finishing our proof.

## Result

2 of 2

We suppose to the contrary, that maximal entry was not on the diagonal, and show contradiction with positive definiteness of that matrix. Click for more details.

## 2. a

Yes. First note that by rank-nullity theorem we have that an  $n \times n$  matrix has an inverse if and only if its rank is  $n$  (this is straightforward consequence of the rank-nullity theorem and the fact that invertibility of the matrix is equivalent with invertibility, i.e. bijectiveness, of its corresponding linear mapping).

Now, in order to relate rank of matrices to the rank of their product we need to prove a lemma which states that for any  $n \times n$  matrix  $B$  and for any  $n \times n$  matrix  $Q$  of rank  $n$ , we have

$$\text{rank}(B) = \text{rank}(QB) = \text{rank}(BQ). \quad (1)$$

Note however that we easily arrive at equivalence of  $\text{rank}(B) = \text{rank}(QB)$  and  $\text{rank}(B) = \text{rank}(BQ)$  by recalling that the rank of a matrix is equal to the rank of its transpose, and thus if we know say the first of those two equalities for any  $B$  and  $Q$  of the full rank, then

$$\begin{aligned} \text{rank}(BQ) &= \text{rank}(Q^t B^t) \\ &= \text{rank}(B^t) \\ &= \text{rank}(B). \end{aligned}$$

Now fix a basis and let  $\mathcal{B}$  and  $\mathcal{Q}$  be the linear mappings corresponding to matrices  $B$  and  $Q$ , respectively. By rank-nullity theorem it is sufficient to show that  $\ker(\mathcal{B}) = \ker(\mathcal{Q} \circ \mathcal{B})$ . As  $Q$  has full rank then  $\ker(\mathcal{Q}) = \{0\}$  so that  $x \in \ker(\mathcal{Q} \circ \mathcal{B})$ , i.e.  $\mathcal{Q}(\mathcal{B}(x)) = 0$  if and only if  $\mathcal{B}(x) = 0$ , which is equivalent with

$$x \in \ker(\mathcal{B}),$$

which is what we wanted to prove.

Now we combine the results of the first and the second paragraph: let  $P$  (and consequently also  $P^t$ ,  $A$  and  $A'$ ) be  $n \times n$  matrices. As by the first paragraph  $P$  has rank  $n$  (and then so does  $P^t$ ), then by the lemma we have the following sequence of equalities

$$\begin{aligned}\text{rank}(A') &= \text{rank}(P^t A P), \\ \text{rank}(A') &= \text{rank}(P^t (A P)), \\ \text{rank}(A') &= \text{rank}(A P), \\ \text{rank}(A') &= \text{rank}(A),\end{aligned}$$

which is what we needed to prove.

## Result

4 of 4

We use rank-nullity theorem to prove that if  $Q$  is an  $n \times n$  matrix with rank  $n$ , then for any  $n \times n$  matrix  $B$  we have that the rank of  $B$  is equal to both the rank of  $QB$  and the rank of  $BQ$ , and use this to give a proof that rank of  $A'$  and  $A$  are equal.

## Section 3

### 1. a

Suppose  $X^* A X$  is a real number for all column vectors  $X$ . As we have that adjoint of any real number (seen as a  $1 \times 1$  matrix) is equal to that real number, then we have for all  $X$

$$X^* A X = (X^* A X)^* = X^* A^* (X^*)^* = X^* A^* X.$$

Therefore, for all  $X$  we have

$$X^* A X - X^* A^* X = 0$$

where by (left and right) distributivity of matrix multiplication we get

$$X^* (A - A^*) X = 0.$$

As this holds for all  $X$  this implies that  $A - A^* = 0$ , i.e.  $A = A^*$ , so that  $A$  is indeed Hermitian.

## Result

2 of 2

We show that  $A$  is Hermitian by using the fact that the adjoint of any real number is equal to that real number. Click for more details.

### 2. a

Let  $A = (a_{ij})$  be a Hermitian matrix, then denote by  $X$  the matrix whose entry on  $(i, j)$ th position is equal to the real part of  $a_{ij}$ . Likewise, denote by  $Y$  the matrix whose entry on  $(i, j)$ th position is equal to the imaginary part of  $a_{ij}$ . Then by definition we have

$$A = X + iY.$$

Furthermore, as  $A$  is Hermitian matrix, then

$$X^* + (iY)^* = X + iY,$$

where by recalling the property of adjoint matrices which states that if  $a$  is a scalar, then  $(aY)^* = \bar{a}Y^*$ , we get

$$X^* - iY^* = X + iY. \quad (1)$$

Here we can "equate real and imaginary parts" in the same way as we can do it for two complex numbers, for note that (1) is equivalent with

$$X^* - X = i(Y + Y^*),$$

and as the matrix on the left hand side is real matrix, and  $Y$  and  $Y^*$  are also real, we must have that

$$X^* = X, \text{ and } Y^* = -Y \quad (2)$$

But note that since  $X$  and  $Y$  are real then  $X^* = X^t$ , so that (2) means that

$$X^t = X, \text{ and } Y^t = -Y,$$

i.e.  $X$  is symmetric and  $Y$  is skew-symmetric.

Now fix a basis of a complex vector space  $V$  and let  $\langle \cdot, \cdot \rangle$  be a positive definite Hermitian form on it, then there exists a Hermitian matrix  $A$  such that for any vectors  $v$  and  $w$  having coordinate vectors  $C$  and  $D$ , respectively, we have

$$\langle v, w \rangle = C^*AD.$$

Now suppose we make  $V$  a real vector space by restricting scalars to real numbers, then coordinate vectors of  $v$  and  $w$  are real matrices, so that with a fixed (real) basis  $C$  and  $D$  are real matrices. Then  $C^*AD = C^tAD$  and

$$\langle v, w \rangle = C^t(X + iY)D = C^tXD + i(C^tYD),$$

where since  $C^tXD$  and  $C^tYD$  are real numbers this is precisely the decomposition outlined in the exercise. Now we just need to recall a proposition which says that  $C^tZD$  is a (skew-)symmetric form if and only if  $Z$  is a (skew-)symmetric matrix to conclude this proof.

## Result

3 of 3

We proceed by decomposing an arbitrary Hermitian matrix into a real part and imaginary part, and then show that real part is a symmetric matrix, while imaginary part is a skew-symmetric matrix. After we have done that the result soon follows by an application of a proposition which states that a form is (skew-)symmetric if and only if its corresponding matrix is (skew-)symmetric.

### 3. a



Denote by  $X_{pq}$  an  $n \times n$  matrix with  $a_{pq} = a_{qp} = 1$  and all other entries 0, and by  $Y_{pq}$  an  $n \times n$  matrix with  $a_{pq} = i$  and  $a_{qp} = -i$ , and all other entries 0. In the previous exercise we showed that any Hermitian matrix  $A$  can be decomposed as

$$A = X + iY$$

where  $X$  and  $Y$  are real matrices with  $X$  being symmetric and  $Y$  skew-symmetric. Furthermore, it is easy to see that this decomposition is unique, for if  $A' = X' + iY'$  then  $X - X' = i(Y - Y')$  and as  $X, X', Y$  and  $Y'$  are real matrices then  $X = X'$  and  $Y = Y'$ .

## Step 2

2 of 6

We now see that is sufficient to show that set  $\{X_{pq} : 1 \leq p \leq q \leq n\}$  is a basis for the (real) vector (sub)space of  $n \times n$  real symmetric matrices, while  $\{Y_{pq} : 1 \leq p < q \leq n\}$  is a basis for the (real) vector (sub)space of all  $n \times n$  real skew-symmetric matrices multiplied by  $i$  -- note that it is straightforward to check that all the matrices  $X_{pq}$  are symmetric and  $Y_{pq}$  are skew-symmetric and that both of these really do form a real vector space.

Let us first check that  $\mathcal{X} = \{X_{pq} : 1 \leq i \leq j \leq n\}$  is indeed a basis for the vector space of  $n \times n$  real symmetric basis. First, linear independence of  $\mathcal{X}$  is immediate, for a linear combination of the form

$$\alpha_1 X_{p_1 q_1} + \cdots + \alpha_n X_{p_n q_n}$$

still has nonzero entries only on positions  $(p_1, q_1), (q_1, p_1), \dots, (p_n, q_n), (q_n, p_n)$  and thus cannot equal some  $X_{pq}$  with  $(p, q) \neq (p_k, q_k)$  for any  $k$ .

Now let  $D = (d_{pq})$  be a real symmetric matrix. Then, by the definition of  $X_{pq}$  and since  $D$  is symmetric, we have

$$D = \sum_{1 \leq p < q \leq n} d_{pq} X_{pq},$$

which concludes our proof that  $\mathcal{X}$  is a basis for the space of all real symmetric matrices.

## Step 4

4 of 6

Analogous reasoning also shows that  $\mathcal{Y} = \{Y_{pq} : 1 \leq p < q \leq n\}$  is a basis for the space of real skew-symmetric matrices. Proof for independence of  $\mathcal{Y}$  is basically the same, while now suppose  $C$  is a real skew-symmetric matrix. Now write  $iC = (ic_{pq})$  then as  $C$  is skew-symmetric and thus  $a_{pq} = -a_{qp}$ , we have

$$iC = \sum_{1 \leq p < q \leq n} c_{pq} Y_{pq}.$$

To complete our proof note that  $\mathcal{X} \cup \mathcal{Y}$  is linearly independent. This follows from linear independence of  $\mathcal{X}$  and  $\mathcal{Y}$  and the fact that all the entries in all the matrices are real (complex) in  $\mathcal{X}$  ( $\mathcal{Y}$ ). To recap: first we showed that a Hermitian matrix has a unique decomposition in the form  $X + iY$  with  $X$  and  $Y$  real matrices,  $X$  symmetric and  $Y$  skew-symmetric, and then we exhibited a basis for all symmetric matrices and for all skew-symmetric matrices time  $i$ , thus its union (as it is linearly independent) gives us our basis.

## Result

6 of 6

We show that the set given by  $\{X_{pq} : 1 \leq p \leq q \leq n\} \cup \{Y_{pq} : 1 \leq p < q \leq n\}$  where  $X_{pq}$  is an  $n \times n$  matrix with  $a_{pq} = a_{qp} = 1$  and all other entries 0 and  $Y_{pq}$  is an  $n \times n$  matrix with  $a_{pq} = i$  and  $a_{qp} = -i$ , and all other entries 0, is a basis for all  $n \times n$  Hermitian matrices. Click for the detailed proof.



#### 4. a

We first show that it is Hermitian. Note that we don't need the hypothesis that it is invertible for this; for we simply have, by rules for computing with adjoint matrices,

$$(A^*A)^* = A^*(A^*)^* = A^*A.$$

Now suppose  $A$  is invertible matrix, we want to show that, for any nonzero complex column vector  $X$ , we have

$$X^*(A^*A)X > 0. \quad (1)$$

We compute

$$X^*(A^*A)X = (X^*A^*)AX = (AX)^*(AX),$$

Note that as  $A$  is an  $n \times n$  matrix and  $X$  is an  $n \times 1$  matrix, then  $AX$  is an  $n \times 1$  matrix, so that  $AX$  is a column vector which we write as  $AX = (t_v + iq_v)_{i=1, \dots, n}$  for  $t_v$  and  $q_v$ . It was shown in the text (see the computation of  $\langle X, X \rangle$ ) that  $(AX)^*(AX)$  is equal to

$$t_1^2 + q_1^2 + \dots + t_n^2 + q_n^2.$$

Thus  $(AX)^*(AX) = 0$  if and only if  $t_1 = q_1 = \dots = t_n = q_n = 0$  which is true if and only if  $AX = 0$ .

However, if  $A$  is invertible then this implies that  $X = 0$ , which contradicts our requirement that  $X$  be nonzero.

Thus, (1) holds, so that  $A^*A$  is positive definite.

#### Result

2 of 2

$A^*A$  being a Hermitian matrix follows simply by rules for computing with adjoint matrices, while its positive definiteness follows from the computation of  $\langle AX, AX \rangle$  for a nonzero column vector  $X$  showing that it is 0 if and only if  $AX = 0$ , which cannot happen if  $A$  is invertible. Click for more details.

#### 5. a

Let  $A$  be a Hermitian positive definite matrix. As  $A = A^*$  we have the following sequence of equalities

$$\begin{aligned} 0A &= 0 \\ (A^* - A)A &= 0 \\ A^*A - AA &= 0 \\ A^2 &= A^*A. \end{aligned}$$

Thus, using the previous exercise, in order to show that  $A^2$  is Hermitian and positive definite, it is sufficient to show that  $A$  is invertible. Suppose  $A$  was not invertible; then there exist a nonzero  $X$  such that  $AX = 0$  (this is

**Theorem 1.2.21**), but then  $X^tAX = X^t0 = 0$ , which is in contradiction with positive definiteness of  $A$ . Thus  $A^2$  is Hermitian and positive definite.

Let again  $A$  be a Hermitian positive definite matrix. We showed in the previous paragraph that it is invertible, hence we can ask whether  $A^{-1}$  is also Hermitian and positive definite. To show that it indeed is, first observe that as

$$AA^{-1} = I$$

then  $(A^{-1})^*A^* = I$  and since  $A$  is Hermitian

$$(A^{-1})^*A = I,$$

i.e.  $(A^{-1})^*$  is the inverse of  $A$  and thus via uniqueness of inverses,

$$(A^{-1})^* = A^{-1}.$$

Now, in order to show that  $A^{-1}$  is positive definite, note that  $A$  being invertible means that the corresponding linear mapping is bijective, and hence for any column vector  $Y$  there exists a column vector  $X$  such that  $AX = Y$ . Thus we compute

$$\begin{aligned} Y^*A^{-1}Y &= (AX)^*A^{-1}(AX) \\ &= X^*A^*X \\ &= X^*AX > 0 \end{aligned}$$

where the second equality follows from properties of computing with adjoint matrices and the third equality follows from the fact that  $X^*A^*X$  is a real number, and hence it is equal to its adjoint, which is exactly  $X^*AX$ . Hence  $A^{-1}$  is also Hermitian and positive definite.

We use the previous exercise to construct matrices  $A$  and  $B$  which are Hermitian and positive definite but whose product is not Hermitian. Let

$$A' = \begin{bmatrix} 1 & 0 \\ -i & 1 \end{bmatrix}, \quad B' = \begin{bmatrix} 2 & 0 \\ -i & 2 \end{bmatrix}.$$

$A'$  and  $B'$  are easily seen to be invertible as  $\det(A') = 1$  and  $\det(B') = 4$ . We compute and put

$$A = (A')^*A' = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}, \quad B = (B')^*B' = \begin{bmatrix} 5 & 2i \\ -2i & 4 \end{bmatrix}.$$

Then by the previous exercise both  $A$  and  $B$  are Hermitian (which can also be seen by just looking at them) and positive definite, but

$$AB = \begin{bmatrix} 12 & 8i \\ -7i & 6 \end{bmatrix},$$

which is obviously not Hermitian, as

$$(AB)^* = \begin{bmatrix} 12 & 7i \\ -8i & 6 \end{bmatrix}$$

and thus  $(AB)^* \neq AB$ .

We show that if  $A$  and  $B$  are Hermitian positive definite, then  $A + B$  is again Hermitian positive definite. That it is Hermitian follows immediately from the rules of computing with adjoint matrices and the fact that  $A$  and  $B$  are Hermitian, for we have

$$(A + B)^* = A^* + B^* = A + B.$$

As they're both positive definite, then for any column vector  $X$  we have

$$X^*AX > 0, \text{ and } X^*BX > 0,$$

but then also, by left and right distributivity of matrix multiplication over addition,

$$\begin{aligned} X^*AX + X^*BX &> 0 \\ X^*(A + B)X &> 0, \end{aligned}$$

which completes the proof that  $A + B$  is positive definite Hermitian.

## Result

5 of 5

We show that  $A^2$ ,  $A^{-1}$  and  $A + B$  are positive definite Hermitian, while  $AB$  is not necessarily Hermitian. Click for the detailed proof.

## 6. a

Since transposition fixes the diagonal, it is easy to see that elements on the diagonal of a Hermitian matrix must be real numbers. This means that an arbitrary Hermitian matrix  $A$  has the form

$$A = \begin{bmatrix} x & z \\ \bar{z} & y \end{bmatrix}$$

for some real numbers  $x$  and  $y$  and a complex number  $z$ .

Recall the the characteristic polynoial of  $A$  is defined as  $p_A(t) = \det(tI - A)$  and eigenvalues of  $A$  are its roots. Thus we have to prove that all the roots of  $p_A$  are real. We compute it as

$$\begin{aligned} p_A(t) &= \det \left( t \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} x & z \\ \bar{z} & y \end{bmatrix} \right) \\ &= \det \left( \begin{bmatrix} t-x & -z \\ -\bar{z} & t-y \end{bmatrix} \right) \\ &= (t-x)(t-y) + z\bar{z} \\ &= t^2 - (x+y)t + xy + z\bar{z}. \end{aligned}$$

Now recall that if  $z = a + bi$  for real  $a, b$ , then  $z\bar{z} = a^2 + b^2$  and hence  $z\bar{z}$  is a real nonnegative number.

Denoting  $c = -x - y$  and  $d = xy + z\bar{z}$  we see that the roots of  $p_A(t)$  are given by

$$t_1 = \frac{-c + \sqrt{c^2 - 4d}}{2}, \quad t_2 = \frac{-c - \sqrt{c^2 - 4d}}{2},$$

so that  $p_A$  has real roots if and only if  $c^2 - 4d \geq 0$ . But note that we have

$$\begin{aligned} (-x - y)^2 - 4(xy + z\bar{z}) &= x^2 + 2xy + y^2 - 4xy + 4z\bar{z} \\ &= (x - y)^2 + 4z\bar{z} \geq 0 \end{aligned}$$

where the last inequality follows from nonnegativity of squares and nonnegativity of  $z\bar{z}$ , finishing our proof.

## Result

2 of 2

We note that a  $2 \times 2$  Hermitian matrix must have a particular form, then we compute its characteristic root directly and show that it must have real roots. Click for the detailed proof.

# Section 4

## 1. a

Suppose  $B$  is an invertible matrix whose columns are orthogonal with respect to the bilinear form  $\langle \cdot, \cdot \rangle$  given by

$$\langle v, w \rangle = X^* A Y$$

for some matrix  $A$  and for column vectors  $X$  and  $Y$  corresponding to vectors  $v$  and  $w$ , respectively.

We write

$$B = [b_1 \quad b_2 \quad \cdots \quad b_n]$$

for some column vectors  $b_1, \dots, b_n$ .

$$b_i = \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ni} \end{bmatrix}, \text{ for } i = 1, \dots, n.$$

By hypothesis we have that  $\langle b_i, b_j \rangle = 0$  for  $i \neq j$ , i.e.

$$\sum_{k,l} b_{ki} a_{kl} b_{lj} = 0,$$

where we take the sum over all  $k, l$  between 1 and  $n$ , as we also take in all the sums in this solution.

Now we compute,

$$\begin{aligned}
 B^*AB &= \begin{bmatrix} \overline{b_1^t} \\ \overline{b_2^t} \\ \vdots \\ \overline{b_n^t} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix} \\
 &= \begin{bmatrix} \langle b_1, b_1 \rangle & \langle b_1, b_2 \rangle & \dots & \langle b_1, b_n \rangle \\ \langle b_2, b_1 \rangle & \langle b_2, b_2 \rangle & \dots & \langle b_2, b_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle b_n, b_1 \rangle & \langle b_n, b_2 \rangle & \dots & \langle b_n, b_n \rangle \end{bmatrix} \\
 &= \begin{bmatrix} \langle b_1, b_1 \rangle & 0 & \dots & 0 \\ 0 & \langle b_2, b_2 \rangle & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \langle b_n, b_n \rangle \end{bmatrix},
 \end{aligned}$$

i.e.  $B^*AB$  is a diagonal matrix with elements  $\langle b_1, b_1 \rangle, \dots, \langle b_n, b_n \rangle$  on the diagonal; assuming that the form is nondegenerate symmetric form on a real vector space or a Hermitian form on a complex vector space, we can indeed show that they are all nonzero. For in that case we have (by **Proposition 8.4.4**) that if the form is nondegenerate then  $A$  is invertible, so that if some  $\langle b_i, b_i \rangle = 0$ , then rank of  $D = B^*AB$  is less than  $n$  and hence it is noninvertible. But then

$$A = (B^*)^{-1}DB^{-1},$$

but this means that  $A$  is noninvertible, as a product of an invertible and a noninvertible matrix is noninvertible. Thus all  $\langle b_i, b_i \rangle \neq 0$  and hence  $D$  is invertible; its inverse is given by

$$D^{-1} = \begin{bmatrix} \frac{1}{\langle b_1, b_1 \rangle} & 0 & \dots & 0 \\ 0 & \frac{1}{\langle b_2, b_2 \rangle} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\langle b_n, b_n \rangle} \end{bmatrix}$$

and hence since

$$(D^{-1}B^*A)B = I$$

by the uniqueness of inverses we get that  $D^{-1}B^*A$  is the inverse of  $B$ .

**Remark.** If  $\langle \cdot, \cdot \rangle$  is the ordinary dot product on the real or complex vector space and if in addition to that the vectors are not only orthogonal but also orthonormal, meaning that also  $\langle b_i, b_i \rangle = 1$  for all  $i$ , then the inverse of  $B$  is its transpose or its adjoint, respectively.

## Result

4 of 4

In the case of the ordinary dot product and orthonormal vector the inverse is just its transpose or its adjoint, depending on whether we're in a real or a complex vector space. In the case of a more general nondegenerate symmetric form on real space or nondegenerate Hermitian form on complex space, we also compute its inverse.

[Click for more details.](#)

2. a



So  $V$  is a real vector space such that  $\langle \cdot, \cdot \rangle$  is a bilinear form on  $V$ . Suppose  $v \in V$ , such that  $\langle v, v \rangle \neq 0$ . Then  $V = \langle v \rangle \oplus v^\perp$ . We have to find the formula for orthogonal projection. Suppose  $u \in V$ . Then  $u$  can be written uniquely as  $\alpha v + u'$ , where  $\alpha \in \mathbb{R}$ , and  $u' \in v^\perp$ . Then, the orthogonal projection of  $u$  is  $u'$ . We know, have to just find the explicit formula for  $u'$ . Observe  $u' = u - \alpha v$ , and  $\langle u', v \rangle = 0$ . Therefore, we have  $\langle u - \alpha v, v \rangle = 0 \implies \alpha = \frac{\langle u, v \rangle}{\langle v, v \rangle}$ . So, we have  $u' = u - \frac{\langle u, v \rangle}{\langle v, v \rangle} v$ . Therefore the orthogonal projection of  $u$  along the vector  $v$  is given by  $u - \frac{\langle u, v \rangle}{\langle v, v \rangle} v$ .

## Result

2 of 2

We just use the fact that  $V = \langle v \rangle \oplus v^\perp$ . Then, one, can deduce the formula just from the definition of orthogonality. See the solution for more details.

## 3. a

Note that  $B$  is an  $n \times n$  matrix, and let  $X$  be a  $n \times 1$  matrix. Then  $AX$  is an  $m \times 1$  matrix, so that

$$\begin{aligned} X^t B X &= X^t A^t A X \\ &= (AX)^t A X \\ &= \sum_{i=1}^m y_i^2 \geq 0, \end{aligned} \tag{1}$$

where

$$AX = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Thus  $B$  is positive semidefinite. Now, in order to prove that  $\text{rank}(A^t A) = \text{rank}(A)$  note that by rank-nullity theorem it is sufficient to show that

$$\dim(\ker A^t A) = \dim(\ker A). \tag{2}$$

Let  $X$  be an  $n \times 1$  matrix (i.e. a column vector of an  $n$  dimensional vector), then note that  $A^t A$  has dimension  $m \times m$  and  $A$  has dimension  $m \times n$ , so that (2) holds if we prove that

$$A^t A X = 0 \text{ if and only if } A X = 0.$$

Suppose first that  $A^t A X = 0$ , then also

$$\begin{aligned} X^t A^t A X &= 0, \\ (AX)^t A X &= 0, \end{aligned}$$

but by (1) this equality only holds if  $y_1 = y_2 = \dots = y_m = 0$ , i.e. if  $AX = 0$ .

Conversely, if  $AX = 0$  then by associativity of matrix multiplication we have

$$A^t A X = A^t (A X) = A^t 0 = 0.$$

Now as  $\ker(A^t A) = \ker(A)$  then their dimension are also the same, and hence by rank-nullity theorem their rank is also the same, and hence  $\text{rank}(A^t A) = \text{rank}(B)$ .

## Result

2 of 2

In order to prove that  $B$  is positive semidefinite we show that  $X^t B X$  is a sum of squares of real numbers, while in order to show that its rank is equal to  $A$  we use rank-nullity theorem to reduce the problem to showing that kernels of  $B$  and  $A$  are the same. Click for more details.



4. a

In order for  $X$  and  $Y$  to be orthogonal we must have

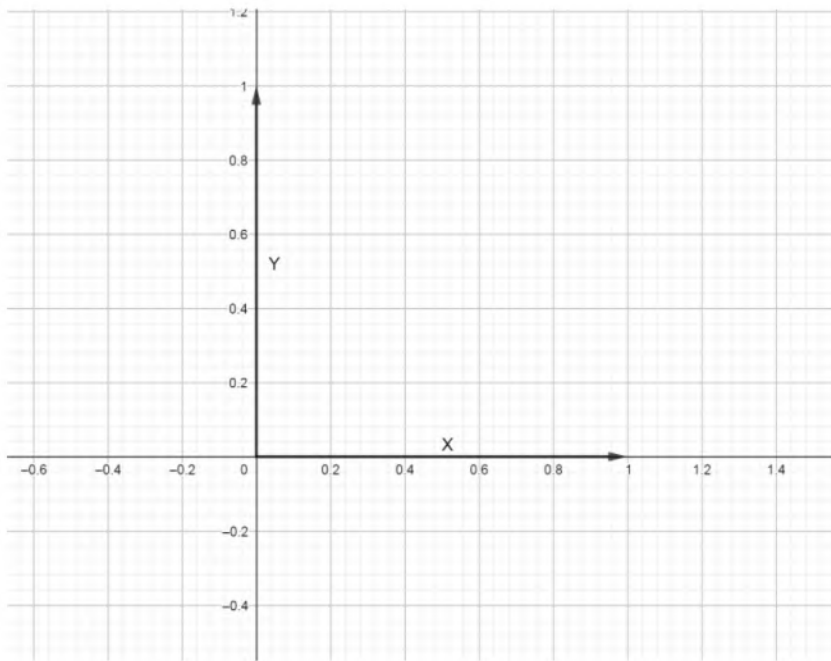
$$x_1y_1 - x_2y_2 = 0.$$

We sketch a few examples using Geogebra. If

$$X = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

then we must have  $y_1 = 0$  and  $y_2 \in \mathbb{R}$ . For example the vector given by

$$Y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



If

$$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

then we must have

$$Y = \begin{bmatrix} y \\ y \end{bmatrix}$$

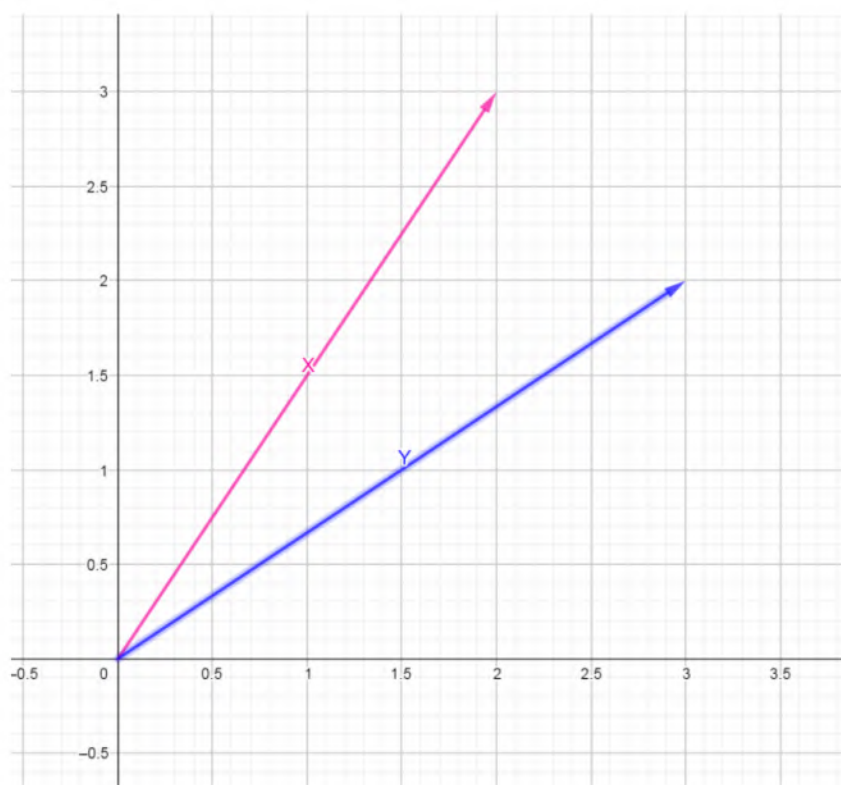
for  $y \in \mathbb{R}$ , say  $y = -1$ .



If

$$X = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

then we must have  $y_1 = \frac{3}{2}y_2$ , e.g.  $y_2 = 2$  and  $y_1 = 3$ .



## Result

We show examples of vectors orthogonal to

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

,

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

. Click for more details.

5. a

(a)

Note that we have, for

$$X = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

and

$$Y = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

, then  $\langle X, Y \rangle = 0$  if and only if

$$\begin{aligned} \begin{bmatrix} x_1 & y_1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} &= 0, \\ \begin{bmatrix} (x_1 + y_1) & (x_1 + y_1) \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} &= 0, \\ (x_1 + y_1)(x_2 + y_2) &= 0 \end{aligned}$$

Thus for choice of  $x_1 = 1, y_1 = 0$  and  $x_2 = 1$  and  $y_2 = -1$  we get two orthogonal vectors, and it is easy to check that they're a basis as we have

$$X - Y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

, i.e. the standard basis is in the span of  $\{X, Y\}$ , so as they're two vectors which span  $\mathbb{R}^2$ , then they're a basis.

(b)

Similarly as in (a) we must find a basis  $X, Y, Z$  such that

$$\begin{aligned} \begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} &= 0, \\ \begin{bmatrix} x_1 + z_1 & 2y_1 + z_1 & x_1 + y_1 + z_1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} &= 0, \\ (x_1 + z_1)x_2 + (2y_1 + z_1)y_2 + (x_1 + y_1 + z_1)z_2 &= 0, \end{aligned} \tag{1}$$

and similarly

$$(x_1 + z_1)x_3 + (2y_1 + z_1)y_3 + (x_1 + y_1 + z_1)z_3 = 0, \tag{2}$$

and

$$(x_2 + z_2)x_3 + (2y_2 + z_2)y_3 + (x_2 + y_2 + z_2)z_3 = 0, \tag{3}$$

each corresponding to  $\langle X, Y \rangle = 0$ ,  $\langle X, Z \rangle = 0$  and  $\langle Y, Z \rangle = 0$ , respectively -- note that for any  $B$  and  $C$  we have  $\langle B, C \rangle = 0$  if and only if  $\langle C, B \rangle = 0$  because  $A$  is symmetric matrix.

First pick  $x_3 = 1$ ,  $y_3 = 0$  and  $z_3 = 0$ , then (2) transforms into  $x_1 + z_1 = 0$  and (3) transforms into  $z_2 = -x_2$ . Then (1) transforms into

$$(2y_1 - x_1)y_2 - y_1x_2 = 0.$$

Thus we're looking for independent vectors of the form

$$\begin{bmatrix} x_1 \\ y_1 \\ -x_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \\ -x_2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

subject to

$$2y_1y_2 - x_1y_2 - y_1x_2 = 0.$$

Putting  $x_1 = 0$  and  $y_1 = 1$  we get

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \\ -x_2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

subject to

$$2y_2 - x_2 = 0,$$

where we put  $x_2 = 2$  and  $y_2 = 1$ , by which we obtained

$$X = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, Y = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, Z = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

As

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{-1}{2}(Y - X - 2Z)$$

then we have that  $X, Y, Z$  are a basis by the analogous reasoning as in **(a)**, and they're orthogonal by construction.

## Result

4 of 4

In **(a)** we find that

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

form an orthogonal basis, while in **(b)** we find that

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

form an orthogonal basis.

6. a

Let

$$e_1^t = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

,

$$e_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$$

and

$$e_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$$

. Then it is not hard to see that  $\{X_1, e_1, e_2, e_3\}$  is a basis for  $\mathbb{R}^4$ , for it consists of four vectors and since

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} = 2X_1 - e_1 + e_2 - e_3,$$

we see that it spans  $\mathbb{R}^4$ .

We now follow the Gram-Schmidt procedure inductively in order to obtain an orthonormal basis. Note that  $\langle X_1, X_1 \rangle = 1$ , and so  $(X_1)$  is an orthonormal basis for the subspace  $V_1$  spanned by  $X_1$ . Now we note that we have a projection  $\pi_1 : V \rightarrow V_1$ , which by the projection formula is given by

$$\pi_1(v) = \frac{\langle X_1, v \rangle}{\langle X_1, X_1 \rangle} X_1,$$

so that in order to obtain the second vector in our basis we compute

$$\begin{aligned} x_2 &= e_1 - \pi_1(e_1) \\ &= e_1 - \frac{\langle X_1, e_1 \rangle}{\langle X_1, X_1 \rangle} X_1 \\ &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 3/4 \\ 1/4 \\ -1/4 \\ -1/4 \end{bmatrix} \end{aligned}$$

where we observe that  $\sqrt{\langle x_2, x_2 \rangle} = \sqrt{3}/2$ , so that the second vector in our orthonormal basis is given by

$$X_2 = \frac{2}{\sqrt{3}} \begin{bmatrix} 3/4 \\ 1/4 \\ -1/4 \\ -1/4 \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 \\ \sqrt{3}/6 \\ -\sqrt{3}/6 \\ -\sqrt{3}/6 \end{bmatrix}.$$

Now we repeat the process, taking  $X_1, X_2$  as an orthonormal basis of subspace  $V_2$ . We again have a projection  $\pi_2 : V \rightarrow V_2$  given by

$$\pi_2(v) = \frac{\langle X_1, v \rangle}{\langle X_1, X_1 \rangle} X_1 + \frac{\langle X_2, v \rangle}{\langle X_2, X_2 \rangle} X_2.$$

Hence we compute

$$\begin{aligned} x_3 &= e_2 - \langle X_1, e_2 \rangle X_1 - \langle X_2, e_2 \rangle X_2 \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \left(-\frac{1}{2}\right) \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} - \frac{\sqrt{3}}{6} \left( \frac{2}{\sqrt{3}} \begin{bmatrix} 3/4 \\ 1/4 \\ -1/4 \\ -1/4 \end{bmatrix} \right) \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/4 \\ -1/4 \\ 1/4 \\ 1/4 \end{bmatrix} - \begin{bmatrix} 3/12 \\ 1/12 \\ -1/12 \\ -1/12 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 2/3 \\ 1/3 \\ 1/3 \end{bmatrix} \end{aligned}$$

where we compute  $\sqrt{\langle x_3, x_3 \rangle} = \sqrt{6}/3$  and hence

$$X_3 = \frac{3}{\sqrt{6}} \begin{bmatrix} 0 \\ 2/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{6}/3 \\ \sqrt{6}/6 \\ \sqrt{6}/6 \end{bmatrix}.$$



Finally we compute the last vector. By completely analogous reasoning as before we compute

$$\begin{aligned} x_4 &= e_3 - \langle X_1, e_3 \rangle X_1 - \langle X_2, e_3 \rangle X_2 - \langle X_3, e_3 \rangle X_3 \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} - \left( -\frac{\sqrt{3}}{6} \right) \begin{bmatrix} \sqrt{3}/2 \\ \sqrt{3}/6 \\ -\sqrt{3}/6 \\ -\sqrt{3}/6 \end{bmatrix} - \frac{\sqrt{6}}{6} \begin{bmatrix} 0 \\ \sqrt{6}/3 \\ \sqrt{6}/6 \\ \sqrt{6}/6 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 1/2 \\ -1/2 \end{bmatrix} \end{aligned}$$

which we normalize in order to obtain

$$X_4 = \begin{bmatrix} 0 \\ 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}.$$

## Result

We use Gram-Schmidt procedure to find that

$$X_1, \begin{bmatrix} \sqrt{3}/2 \\ \sqrt{3}/6 \\ -\sqrt{3}/6 \\ -\sqrt{3}/6 \end{bmatrix}, \begin{bmatrix} 0 \\ \sqrt{6}/3 \\ \sqrt{6}/6 \\ \sqrt{6}/6 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

form an orthonormal basis for  $\mathbb{R}^4$ . [Click to see more details.](#)

7. a

Let

$$u_1^t = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$$

,

$$u_2^t = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$

, and

$$u_3^t = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$$

. First we normalize  $u_1$  so that its dot product with itself is equal to 1, i.e. divide each of its elements by  $\sqrt{\langle u_1, u_1 \rangle} = \sqrt{2}$ . Hence the first vector in our basis is

$$v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}.$$

Now we have the projection  $\pi_1$  of  $V$  to  $V_1$ , the subspace spanned by  $v_1$ , given by

$$\pi_1(v) = v_1 \frac{\langle v_1, v \rangle}{\langle v_1, v_1 \rangle}.$$

Hence we compute

$$u = u_2 - \pi_1(u_2) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{\sqrt{2}}{2} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix},$$

which we divide by  $\sqrt{\langle u, u \rangle} = \sqrt{6}/2$  in order to obtain

$$v_2 = \begin{bmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} = \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}.$$

Finally we find the third vector in our basis, again forming a projection  $\pi_2$  from  $V$  to  $V_2$ , subspace spanned by  $v_1$  and  $v_2$ , given by

$$\pi_2(v) = v_1 \frac{\langle v_1, v \rangle}{\langle v_1, v_1 \rangle} + v_2 \frac{\langle v_2, v \rangle}{\langle v_2, v_2 \rangle}.$$

Hence the vector we have to normalize is

$$\begin{aligned} t &= u_3 - \pi_2(u_3) \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{\sqrt{2}}{2} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} - \frac{\sqrt{6}}{6} \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/6 \\ -1/6 \\ 1/3 \end{bmatrix} \\ &= \begin{bmatrix} -2/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \end{aligned}$$

where by normalizing it by dividing it by  $\sqrt{\langle t, t \rangle} = 2/\sqrt{3}$  we get

$$v_3 = \begin{bmatrix} -\sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \end{bmatrix}.$$

## Result

By using the Gram-Schmidt procedure we arrive at the following orthonormal basis:

$$\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}, \begin{bmatrix} -\sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \end{bmatrix}.$$

[Click for more details.](#)

8. a

As  $A$  is symmetric and by the characterization of positive definite symmetric matrices in relation to their minors (see **Theorem 8.4.19**) we can easily see that  $A$  is positive definite. Now we can use Gram-Schmidt procedure to find an orthonormal basis.

Now, we have that the vectors  $P = (1, 2)^t$  and  $Q = (2, 1)^t$  form a basis of  $\mathbb{R}^2$ . First we normalize  $P$ ; in order to do that we compute

$$\langle P, P \rangle = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 14,$$

so that we have to divide  $P$  by  $\sqrt{14}$ , so that the first vector in our basis is

$$v_1 = \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \end{bmatrix}.$$

Furthermore, we have a projection  $\pi$  from  $\mathbb{R}^2$  to the subspace spanned by  $v_1$  is given by

$$\pi(v) = \frac{\langle v_1, v \rangle}{\langle v_1, v_1 \rangle} v.$$

Then

$$\pi(Q) = \langle v_1, Q \rangle v = \frac{13}{\sqrt{14}} \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \end{bmatrix} = \begin{bmatrix} 13/14 \\ 26/14 \end{bmatrix},$$

so that

$$u_2 = Q - \pi(Q) = \begin{bmatrix} 15/14 \\ -12/14 \end{bmatrix}$$

which we normalize by dividing it with square root of

$$\langle u_2, u_2 \rangle = \begin{bmatrix} 15/14 & -12/14 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 15/14 \\ -12/14 \end{bmatrix} = \frac{27}{14},$$

so that the second vector of our basis is

$$v_2 = \begin{bmatrix} \frac{15\sqrt{14}}{14\sqrt{27}} \\ \frac{-12\sqrt{14}}{14\sqrt{27}} \end{bmatrix}$$

## Result

We use Gram-Schmidt procedure in order to find that

$$\begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \end{bmatrix}$$

and

$$\begin{bmatrix} \frac{15\sqrt{14}}{14\sqrt{27}} \\ \frac{-12\sqrt{14}}{14\sqrt{27}} \end{bmatrix}$$

form an orthonormal basis.

9. a

Note that a basis of all real polynomials of degree at most 2 is given by  $1, x, x^2$ . Now we can apply Gram-Schmidt procedure in order to produce an orthonormal basis.

First note that

$$\int_{-1}^1 1 dx = 2,$$

hence our first vector is (by linearity of the integral)

$$v_1 = \frac{1}{2}.$$

Now forming a projection  $\pi_1$  from  $P$  to  $P_0$ , the subspace generated by  $v_1$ , we have by projection formula that it's given by

$$\pi_1(f) = \frac{\langle v_1, f \rangle}{\langle v_1, v_1 \rangle} v_1.$$

Hence

$$\pi_1(x) = \frac{1}{2} \int_{-1}^1 x dx = 0$$

and hence our second vector is given by

$$v_2 = \frac{x - \pi_1(x)}{\sqrt{\langle x - \pi_1(x), x - \pi_1(x) \rangle}} = \frac{x}{\sqrt{\int_{-1}^1 x^2 dx}} = \frac{x}{\sqrt{3/2}} = \frac{\sqrt{2}x}{\sqrt{3}}.$$

(Where we divided with  $\sqrt{\langle x - \pi_1(x), x - \pi_1(x) \rangle} = \sqrt{\langle x, x \rangle}$  in order to normalize.)

Now we again form a projection  $\pi_2$  of  $P$  to the subspace  $P_1$ , spanned by the vectors  $v_1$  and  $v_2$ , where we get

$$\begin{aligned} \pi_2(x^2) &= \frac{1}{2} \langle 1/2, x^2 \rangle + \frac{\sqrt{2}x}{\sqrt{3}} \left\langle \frac{\sqrt{2}x}{\sqrt{3}}, x^2 \right\rangle \\ &= \frac{1}{2} \cdot \frac{1}{3} \frac{\sqrt{2}x}{\sqrt{3}} \cdot 0 \\ &= \frac{1}{6}. \end{aligned}$$

Hence the final vector in our orthonormal basis is given by

$$\begin{aligned} v_3 &= \frac{x^2 - 1/6}{\sqrt{\langle x^2 - 1/6, x^2 - 1/6 \rangle}} \\ &= \frac{x^2 - 1/6}{\sqrt{\int_{-1}^1 (x^2 - 1/6)^2 dx}} \\ &= \frac{x^2 - 1/6}{\sqrt{7/30}} \\ &= \frac{\sqrt{30}}{\sqrt{7}} x^2 - \frac{\sqrt{30}}{6\sqrt{7}}. \end{aligned}$$

## Result

We use Gram-Schmidt procedure on the basis  $\{1, x, x^2\}$  in order to obtain an orthonormal basis

$$\left\{ \frac{1}{2}, \frac{\sqrt{2}x}{\sqrt{3}}, \frac{\sqrt{30}}{\sqrt{7}} x^2 - \frac{\sqrt{30}}{6\sqrt{7}} \right\}.$$

[Click for more details.](#)

10. a

Recall that if  $A = (a_{ij})$  is an  $n \times n$  matrix, then

$$\text{trace}(A) = \sum_{i=1}^n a_{ii}.$$

First we check that  $\langle, \rangle$  does indeed define a bilinear form. Let  $A = (a_{ij})$  and  $B = (b_{ij})$  be  $n \times n$  matrices. Then it is straightforward, from the definitions of matrix and scalar multiplication, to show that  $\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B)$  and that  $\text{trace}(aA) = a \text{trace}(A)$  for  $a \in \mathbb{R}$ , as well as that  $\text{trace}(A) = \text{trace}(A^t)$  (since transposition doesn't change the diagonal).

Now we have

$$\begin{aligned} \langle rA, B \rangle &= \text{trace}((rA)^t B) \\ &= \text{trace}(r(A^t B)) \\ &= r \text{trace}(A^t B) \\ &= r \langle A, B \rangle \end{aligned}$$

and completely analogously we could show that  $\langle A, rB \rangle = r \langle A, B \rangle$ . Similarly, if  $C$  is an  $n \times n$  matrix, then

$$\begin{aligned} \langle (A + B), C \rangle &= \text{trace}((A + B)^t C) \\ &= \text{trace}(A^t C + B^t C) \\ &= \text{trace}(A^t C) + \text{trace}(B^t C) \\ &= \langle A, C \rangle + \langle B, C \rangle, \end{aligned}$$

and analogously we could show that  $\langle A, B + C \rangle = \langle A, B \rangle + \langle A, C \rangle$ .

Now let us show it is positive definite. Suppose  $A$  is a nonzero matrix -- i.e. such that at least one  $a_{ij} \neq 0$ . We have that  $A^t A$  is an  $n \times n$  matrix with the entry in  $m$ th row and  $k$ th column being equal to

$$\sum_{t=1}^n a_{tm} a_{tk},$$

and hence the  $k$ th entry on the main diagonal is equal to

$$\sum_{t=1}^n (a_{tk})^2.$$

Now we can compute the trace of  $A^t A$ , so that

$$\begin{aligned} \langle A, A \rangle &= \text{trace}(A^t A) \\ &= \sum_{k=1}^n \sum_{t=1}^n (a_{tk})^2 \\ &\geq (a_{ij})^2 > 0, \end{aligned} \tag{1}$$

i.e. it is a positive definite bilinear form.

Now let  $A = (a_{ij})$  and  $B = (b_{ij})$  be two  $n \times n$  matrices, then we have that the  $(i, j)$ th entry in  $A^t B$  is equal to

$$\sum_{t=1}^n a_{ti} b_{tj},$$

and hence  $i$ th entry on the main diagonal is equal to

$$\sum_{t=1}^n a_{ti} b_{ti},$$

so that the trace of  $A^t B$  is equal to

$$\sum_{i=1}^n \sum_{t=1}^n a_{ti} b_{ti}. \quad (2)$$

Let  $e_{ij}$  be the  $n \times n$  matrix with  $(i, j)$ th entry 1 and all the rest 0, then the matrices  $e_{ij}$  for  $1 \leq i, j \leq n$  form the standard basis of real  $n \times n$  matrices. We claim that this is an orthonormal basis with regards to our form.

First we have to check that for any  $i, j$  we have

$$\langle e_{ij}, e_{ij} \rangle = 1,$$

but note that this follows immediately from our computation of trace of  $A^t A$  (see (1)). In the case of  $e_{ij}$  all summands in this double sum are equal to 0 except the one which corresponds to  $k = i$  and  $t = j$ , which is equal to  $1^2 = 1$ , and hence the whole sum is equal to 1.

Finally we have to check that for  $i, j, p, q$  such that  $(i, j) \neq (p, q)$  we have

$$\langle e_{ij}, e_{pq} \rangle = 0.$$

This follows directly from (2), for if there were any summand in (2) which were nonzero, that would mean that some entry is nonzero in both  $e_{ij}$  and  $e_{pq}$ , but this is not so by definition of those matrices. Hence, the matrices  $e_{ij}$  form an orthonormal basis.

## Result

6 of 6

The fact that the form is bilinear follows from the properties of trace, while its positive definiteness follows from computation showing that the trace of  $A^t A$  is a sum of squares. We also show that the standard basis of  $n \times n$  matrices, which consists of matrices  $e_{ij}$  which have the  $(i, j)$ th entry 1 and all the rest 0, also forms an orthonormal basis in regards to our form.

11. a



(a)

Let us first show that

$$(W_1 + W_2)^\perp \subseteq W_1^\perp \cap W_2^\perp. \quad (1)$$

Suppose  $v \in (W_1 + W_2)^\perp$ . Then  $v$  is orthogonal to every  $w = w_1 + w_2 \in W_1 + W_2$ , i.e.

$$\langle v, w_1 + w_2 \rangle = 0$$

for any choice of  $w_1 \in W_1$  and  $w_2 \in W_2$ . But as  $0$  is in a vector subspace, then we can choose  $w_1 = 0$  in order to get that  $v \in W_2^\perp$ , and  $w_2 = 0$  in order to get  $v \in W_1^\perp$ , and hence  $v$  is in their intersection, showing (1).

Now let us show that

$$W_1^\perp \cap W_2^\perp \subseteq (W_1 + W_2)^\perp. \quad (2)$$

If  $v \in W_1^\perp \cap W_2^\perp$  then we have that  $\langle v, w_1 \rangle = 0$  and  $\langle v, w_2 \rangle = 0$ , for all  $w_1 \in W_1$  and  $w_2 \in W_2$ . But then

$$0 = \langle v, w_1 \rangle + \langle v, w_2 \rangle = \langle v, w_1 + w_2 \rangle$$

by the bilinearity of a bilinear form, and hence as any  $w \in W_1 + W_2$  can be written in the form  $w = w_1 + w_2$  with  $w_1 \in W_1$  and  $w_2 \in W_2$ , we have that  $v \in (W_1 + W_2)^\perp$ , showing (2).

(b)

Let us explain what it means for a vector to be in  $W^{\perp\perp}$ . It means that it is orthogonal to every vector which is orthogonal to every vector in  $W$ .

Written like this the inclusion becomes almost obvious, but let us be slightly more formal and note that  $w \in W^{\perp\perp}$  if and only if  $\langle v, w \rangle = 0$  -- as the form is symmetric we need not worry about the order of the elements here -- for any  $v \in W^\perp$ .

Now let  $w \in W$ , then by the definition of  $W^\perp$  we have that, for any  $v \in W^\perp$ ,  $\langle v, w \rangle = 0$ , and hence we see that any  $w \in W$  satisfies the defining property of  $W^{\perp\perp}$ , showing the inclusion.

### Step 3

3 of 4

Suppose  $W_1 \subset W_2$  and let  $w \in W_2^\perp$ . Then for any  $w_2 \in W_2$ ,

$$\langle w, w_2 \rangle = 0.$$

But since  $W_1 \subset W_2$  then also

$$\langle w, w_1 \rangle = 0$$

for any  $w_1 \in W_1$ , and hence  $w \in W_1^\perp$ , showing the inclusion.

### Result

4 of 4

In (a) part we use the basic properties of bilinear product and subspace sums in order to show the set equality, in

(b) we just write out the definition of  $W^{\perp\perp}$ , while in (c) we use the fact that if  $W_1 \subset W_2$  and for some  $w$  is orthogonal to every element of  $W_2$ , then  $w$  is also orthogonal to every element of  $W_1$ .

12. a

Let  $V = \mathbb{R}^{2 \times 2}$  be the vector space of real  $2 \times 2$  matrices.

[Comment](#)

Step 2 of 5 ^

(a)

To determine the matrix of the bilinear form  $\langle A, B \rangle = \text{trace}(AB)$  on  $V$  with respect to the standard basis  $\{e_{ij}\}$ :

Let  $\langle A, B \rangle_i = \text{trace}(A^i B)$ .

Since,

$$\begin{aligned}\langle e_{ij}, e_{kl} \rangle &= \langle e'_{ij}, e_{kl} \rangle_i \\ &= \langle e_{ji}, e_{kl} \rangle_i\end{aligned}$$

This implies that,

$$\langle e_{ij}, e_{kl} \rangle = \begin{cases} 1 & \text{if } (j, i) = (k, l) \\ 0 & \text{otherwise} \end{cases}$$

It follows that the matrix of the form on  $V = \mathbb{R}^{2 \times 2}$  is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Where the ordered basis is  $(e_{11}, e_{12}, e_{21}, e_{22})$ .

(b)

To determine the signature of this form;

Consider the ordered basis:

$$\left( e_{11}, e_{22}, \frac{(e_{12} + e_{21})}{2}, \frac{(e_{12} - e_{21})}{2} \right)$$

It is to check that this is an orthogonal basis and the matrix with respect to this ordered basis is, in block form,

$$\begin{pmatrix} I_3 & 0 \\ 0 & -I_1 \end{pmatrix}$$

Therefore, the signature is  $\boxed{(3,1)}$ .

[Comments \(1\)](#)

Step 4 of 5 ^

(c)

Find an orthogonal basis for this form.

The orthogonal basis for this form is:

$$\left( e_{11}, e_{22}, \frac{(e_{12} + e_{21})}{2}, \frac{(e_{12} - e_{21})}{2} \right)$$

(d)

Determine the signature of the form  $\text{trace } AB$  on the space  $\mathbb{R}^{n \times n}$  of real  $n \times n$  matrices.

Consider the set:

$$\{e_1, \dots, e_n\} \cup \{q_{ij}, m_{ij}\}_{1 \leq i < j \leq n}$$

The above set is an orthogonal basis, where  $q_{ij} = \frac{(e_{ij} + e_{ji})}{2}$  and  $m_{ij} = \frac{(e_{ij} - e_{ji})}{2}$ .

As in the  $2 \times 2$  case, the matrices  $\{e_{ii}\}$  and the matrices  $\{q_{ij}\}$  and  $\{m_{ij}\}$  are pair with themselves to a value of 1 and  $-1$ .

The number of negative ones is  $\binom{n}{2}$ .

Therefore, the signature is:

$$\left(n^2 - \binom{n}{2}, \binom{n}{2}\right) = \left(\frac{n^2 + n}{2}, \frac{n^2 - n}{2}\right)$$

13. a

(a)

To decide whether or not the rule  $\langle A, B \rangle = \text{trace}(A^* B)$  defines a Hermitian form on the space  $\mathbb{C}^{n \times n}$  of complex matrices, and if so, to determine its signature;

[Comment](#)

Step 2 of 4 ^

First show that  $\langle A, B \rangle = \text{trace}(A^* B)$  defines a Hermitian,

Since  $\text{trace}(A^* B) = \text{trace}(B^* A)$

So,

$$\begin{aligned} \text{trace}\left((A^* B)^*\right) &= \text{trace}\left(B^* (A^*)^*\right) \\ &= \text{trace}(B^* A) \\ &= \text{trace}(A^* B) \end{aligned}$$

Hence, the rule  $\langle A, B \rangle = \text{trace}(A^* B)$  defined a Hermitian form on the space  $\mathbb{C}^{n \times n}$  of complex matrices.

Determine the signature of the form  $\text{trace } AB$  on the space  $\mathbb{C}^{n \times n}$  of complex  $n \times n$  matrices.

Consider the set:

$$\{e_1, \dots, e_n\} \cup \{q_{ij}, m_{ij}\}_{1 \leq i < j \leq n}$$

The above set is an orthogonal basis, where  $q_{ij} = \frac{(e_{ij} + e_{ji})}{2}$  and  $m_{ij} = \frac{(e_{ij} - e_{ji})}{2}$ .

As in the  $2 \times 2$  case, the matrices  $\{e_{ii}\}$  and the matrices  $\{q_{ij}\}$  and  $\{m_{ij}\}$  are pair with themselves to a value of 1 and  $-1$ .

The number of negative ones is  $\binom{n}{2}$ .

Therefore, the signature is:

$$\left(n^2 - \binom{n}{2}, \binom{n}{2}\right) = \left(\frac{n^2 + n}{2}, \frac{n^2 - n}{2}\right)$$

(b)

Defined the rule  $\langle A, B \rangle = \text{trace}(\bar{A}B)$  defines a Hermitian form on the space  $\mathbb{C}^{n \times n}$  of complex matrices, and if so, to determine its signature.

Here,  $\text{trace}(\bar{A}B)$  cannot be defined a Hermitian form on the space  $\mathbb{C}^{n \times n}$  of complex matrices because  $\text{trace}(\bar{A}B) \neq \text{trace}((\bar{A}B)^*)$

$$\begin{aligned} \text{trace}((\bar{A}B)^*) &= \text{trace}(B^* \bar{A}^*) \\ &\neq \text{trace}(\bar{A}B) \end{aligned}$$

Therefore, the rule  $\langle A, B \rangle = \text{trace}(\bar{A}B)$  cannot be defined a Hermitian form on the space  $\mathbb{C}^{n \times n}$  of complex matrices.

14. a

We prove this by induction on the dimension of  $A$ . Note that (matrices representing) row and column operations are invertible, and therefore their multiplication is also going to be invertible; furthermore, if we multiply  $A$  (from the left) by a number of row transformations, call them  $E_1, \dots, E_n$ , and multiply  $A$  (from the right) by matrices corresponding to those same transformations applied to columns, then the resulting matrix is going to equal

$$E_n \cdots E_1 A E_1^t \cdots E_n^t.$$

If it is diagonal, then  $P = E_1^t \cdots E_n^t$  gives us the required  $P$ . Hence, it is sufficient to find a number of transformations which are applied to both rows and columns, which diagonalize  $A$ .

(Base case) Let  $n = 2$ . If  $A$  is a  $2 \times 2$  real symmetric matrix then

$$A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

for some real numbers  $a, b, c$ .

We have three distinct cases to consider. First suppose that  $a \neq 0$ , then (we abbreviate row  $i$  by  $R_i$  and column  $i$  by  $C_i$ )

$$\begin{bmatrix} a & c \\ c & b \end{bmatrix} \xrightarrow{R_2 = aR_2 - cR_1} \begin{bmatrix} a & c \\ 0 & ab - c^2 \end{bmatrix} \xrightarrow{C_2 = aC_2 - cC_1} \begin{bmatrix} a & 0 \\ 0 & a^2b - ac^2 \end{bmatrix}.$$

Analogous operations can also be performed if  $b \neq 0$ . If, on the other hand, we have both  $a = 0$  and  $b = 0$ , then

$$\begin{aligned} \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix} &\xrightarrow{R_2 = R_2 - R_1} \begin{bmatrix} 0 & c \\ c & -c \end{bmatrix} \\ &\xrightarrow{C_2 = C_2 - C_1} \begin{bmatrix} 0 & c \\ c & -2c \end{bmatrix} \\ &\xrightarrow{C_1 = C_1 + \frac{1}{2}C_2} \begin{bmatrix} \frac{c}{2} & c \\ 0 & -2c \end{bmatrix} \\ &\xrightarrow{R_1 = R_1 + \frac{1}{2}R_2} \begin{bmatrix} \frac{c}{2} & 0 \\ 0 & -2c \end{bmatrix}. \end{aligned}$$

**(Step case)**

Suppose now that any  $n \times n$  real symmetric matrix can be transformed to diagonal via row and column operations in the manner described in the first paragraph, and let  $A = (a_{ij})$  be a  $(n+1) \times (n+1)$  real symmetric matrix. Without the loss of generality say  $a_{11} \neq 0$ ; if not, then we can "bring" a nonzero entry to  $a_{11}$  by row and column operations with the matrix remaining symmetric. Then we can use the  $a_{11}$  as pivot element and for each  $i = 2, \dots, n+1$  perform the operation  $R_i = a_{11}R_i - a_{i1}R_1$  and  $C_i = a_{11}C_i - a_{i1}C_1$ ; these operations are as outlined before and they transform our matrix to

$$A = \begin{bmatrix} a_{11} & 0 \\ 0 & A' \end{bmatrix}$$

with  $A'$  being a symmetric  $n \times n$  matrix. Now induction hypothesis finishes our proof, since row and column operations on  $A'$  leave the first row and column as they are, and hence we obtain a diagonalization of  $A$  via row and column operations which give rise to matrix  $P$  as outlined in the first section.

**Remark.** Development of the ideas from this proof leads into an algorithm for finding such a matrix  $P$ . Let  $A$  be an  $n \times n$  symmetric matrix,  $I$   $n \times n$  identity matrix, and if we form a block matrix

$$B = \begin{bmatrix} A & I \end{bmatrix},$$

then by the proof we know we can apply row and column operations to  $B$  so that it is transformed into

$$B' = \begin{bmatrix} D & P \end{bmatrix}$$

with  $D$  diagonal. Then  $P^t A P = D$ .

**Result**

5 of 5

We show how a sequence of row and column operations, if aptly chosen, leads to such a  $P$ . Then we prove, via induction on the dimension of the symmetric matrix in question, that such a sequence of row and column operations can always be found.

15. a

In order to use the projection formula, we first find an orthogonal basis for  $W$ , for which we use the Gram-Schmidt procedure (but without normalizing elements, as we don't need an orthonormal basis). Let

$$v_1 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^t$$

$$u_2 = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$$

. Then  $v_1$  is the first vector in our orthogonal basis. Next we form a projection  $\pi$  from  $\mathbb{R}^3$  to subspace spanned by  $v_1$ , so that

$$\pi(u_2) = \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} v_1,$$

where we compute

$$\langle v_1, u_2 \rangle = 1,$$

$$\langle v_1, v_1 \rangle = 2,$$

so that

$$\pi(u_2) = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 \end{bmatrix},$$



so that the second vector in our orthogonal basis is

$$v_2 = u_2 - \pi(u_2) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}.$$

Now by projection formula, the projection of the vector

$$v = [1 \ 0 \ 0]^t$$

to  $W$  is

$$\begin{aligned} \frac{\langle v_1, v \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle v_2, v \rangle}{\langle v_2, v_2 \rangle} v_2 &= \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{-1/2}{6/4} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/6 \\ -1/6 \\ -1/3 \end{bmatrix} \\ &= \begin{bmatrix} 2/3 \\ 1/3 \\ -1/3 \end{bmatrix}. \end{aligned}$$

## Result

We use the projection formula in order to show that the orthogonal projection of that vector is

$$[2/3 \ 1/3 \ -1/3]^t$$

## 16. a

In order to use projection formula we have to first find an orthogonal basis for the subspace of skew-symmetric matrices. Recall that if a real matrix is skew-symmetric if it's diagonal entries are 0 and we have  $a_{ij} = -a_{ji}$ . Indeed it easy to check that the following matrices form a basis

$$B_1 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

We check whether they form an orthogonal basis. Note that if  $A$  is skew-symmetric then  $A^t = -A$ , and since  $B_1, B_2, B_3$ , are skew-symmetric, we have

$$\langle B_i, B_j \rangle = \text{trace}(B_i^t B_j) = \text{trace}(-(B_i) B_j) = -\text{trace}(B_i B_j), \quad (1)$$

and hence we only need to compute  $\text{trace}(B_i B_j)$  for any  $i$  and  $j$ . We have

$$B_1 B_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B_2 B_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, B_3 B_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (2)$$

and hence  $\langle B_i, B_i \rangle \neq 0$  for  $i = 1, 2, 3$ .



We just need to check that they're orthogonal, for which we compute

$$B_1B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_1B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_2B_3 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

from where we see that  $\langle B_i, B_j \rangle = 0$  for  $i \neq j$ , i.e. it indeed is an orthogonal base.

Now we can projection formula which says that if  $\pi$  is an orthogonal projection from  $V$  to  $W$ , then

$$\pi(A) = \sum_{i=1}^3 \frac{\langle B_i, A \rangle}{\langle B_i, B_i \rangle} B_i.$$

Note that we have already computed  $\langle B_i, B_i \rangle = 2$  in (2), using (1). Now let  $A$  be the matrix given in the exercise,

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 1 & 3 & 0 \end{bmatrix},$$

then

$$B_1A = \begin{bmatrix} 0 & 0 & 1 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_2A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 0 \\ -1 & -2 & 0 \end{bmatrix}, \quad B_3A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

and hence using (1)

$$\langle B_1A \rangle = 2, \quad \langle B_2A \rangle = -1, \quad \langle B_3A \rangle = -2,$$

so that

$$\pi(A) = B_1 - 2B_2 - B_3 = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}.$$

## Result

3 of 3

We first find an orthogonal base for the subspace of all skew-symmetric subspaces and then use projection formula to obtain that the projection is

$$\begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

17. a

As per the method outlined in the previous chapter (3.5.13) we have that basechange matrix from the standard basis to  $B$  is

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Hence the representation of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^t$$

in basis  $B$  is

$$P^{-1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 2 & 2 & 2 \\ 3 & -3 & 0 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1/3 + x_2/3 + x_3/3 \\ x_1/2 - x_2/2 \\ x_1/6 + x_2/6 - x_3/3 \end{bmatrix}$$

which we can see is the same result that was obtained in the **Example 8.4.14** through the use of projection formula.

## Result

2 of 2

We compute it and find that the result is the same as with using the projection formula. Click for more details.

## 18. a

### Projection formula:

Let  $\langle \cdot, \cdot \rangle$  be a symmetric form on a real vector space  $V$  or a Hermitian form on a complex vector space  $V$ , and let  $W$  be a subspace of  $V$  on which the form is nondegenerate. If  $(w_1, \dots, w_n)$  is an orthogonal basis for  $W$ , the orthogonal projection  $\pi: V \rightarrow W$  is given by the formula  $\pi(v) = w_1 c_1 + \dots + w_n c_n$ ,

Where,

$$c_i = \frac{\langle w_i, v \rangle}{\langle w_i, w_i \rangle}$$

To find the matrix of a projection  $\pi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$  such that the image of the standard bases of  $\mathbb{R}^3$  forms an equilateral triangle and  $\pi(e_1)$  points in the direction of the  $x$ -axis;

Standard basis as shown below,

$$e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$$

Then, the vector  $v = e_1 + e_2 + e_3$  where  $v \in \mathbb{R}^3$

$$\begin{aligned} v &= (1, 0, 0) + (0, 1, 0) + (0, 0, 1) \\ &= (1, 1, 1) \end{aligned}$$

Apply the projection formula,

$$\begin{aligned} \pi(v) &= e_1 c_1 + e_2 c_2 + e_3 c_3 \\ &= (1, 0, 0) c_1 + (0, 1, 0) c_2 + (0, 0, 1) c_3 \end{aligned}$$

Now find  $c_1, c_2$  and  $c_3$  as shown below,

$$\begin{aligned} c_1 &= \frac{\langle e_1, v \rangle}{\langle e_1, e_1 \rangle} \\ &= \frac{\langle (1, 0, 0), (1, 1, 1) \rangle}{\langle (1, 0, 0), (1, 0, 0) \rangle} \\ &= \frac{1}{1} \\ &= 1 \end{aligned}$$

$$\begin{aligned}
 c_2 &= \frac{\langle e_2, v \rangle}{\langle e_2, e_2 \rangle} \\
 &= \frac{\langle (0, 1, 0), (1, 1, 1) \rangle}{\langle (0, 1, 0), (0, 1, 0) \rangle} \\
 &= \frac{1}{1} \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 c_3 &= \frac{\langle e_3, v \rangle}{\langle e_3, e_3 \rangle} \\
 &= \frac{\langle (0, 0, 1), (1, 1, 1) \rangle}{\langle (0, 0, 1), (0, 0, 1) \rangle} \\
 &= \frac{1}{1} \\
 &= 1
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 \pi(v) &= e_1 c_1 + e_2 c_2 + e_3 c_3 \\
 &= (1, 0, 0) \cdot 1 + (0, 1, 0) \cdot 1 + (0, 0, 1) \cdot 1 \\
 &= (1, 1, 1) \\
 &= v
 \end{aligned}$$

Since the image of the standard bases of  $\mathbb{R}^3$  belongs to  $\mathbb{R}^2$ , so

$$\pi(v) = (1, 0)$$

Therefore, matrix of a projection  $\pi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$  such that the image of the standard bases of  $\mathbb{R}^3$  forms an equilateral triangle is  $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ .

Now find  $\pi(e_1)$  points in the direction of the  $x$ -axis.

Apply orthogonal projection,

$$\begin{aligned}
 \pi(e_1) &= e_1 c_1 + e_2 c_2 + e_3 c_3 \\
 &= (1, 0, 0) c_1 + (0, 1, 0) c_2 + (0, 0, 1) c_3
 \end{aligned}$$

Now find  $c_1, c_2$  and  $c_3$  as shown below,

$$\begin{aligned}
 c_1 &= \frac{\langle e_1, e_1 \rangle}{\langle e_1, e_1 \rangle} \\
 &= \frac{\langle (1, 0, 0), (1, 0, 0) \rangle}{\langle (1, 0, 0), (1, 0, 0) \rangle} \\
 &= \frac{1}{1} \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 c_2 &= \frac{\langle e_2, e_1 \rangle}{\langle e_2, e_2 \rangle} \\
 &= \frac{\langle (0, 1, 0), (1, 0, 0) \rangle}{\langle (0, 1, 0), (0, 1, 0) \rangle} \\
 &= 0
 \end{aligned}$$

Find  $c_3$ ;

$$\begin{aligned} c_3 &= \frac{\langle e_3, e_1 \rangle}{\langle e_3, e_3 \rangle} \\ &= \frac{\langle (0, 0, 1), (1, 0, 0) \rangle}{\langle (0, 0, 1), (0, 0, 1) \rangle} \\ &= 0 \end{aligned}$$

Therefore,

$$\begin{aligned} \pi(e_1) &= e_1 c_1 + e_2 c_2 + e_3 c_3 \\ &= (1, 0, 0) \cdot 1 + (0, 1, 0) \cdot 0 + (0, 0, 1) \cdot 0 \\ &= (1, 0, 0) \\ &= e_1 \end{aligned}$$

Since the image of the standard bases of  $\mathbb{R}^3$  belongs to  $\mathbb{R}^2$ , so

$$\pi(e_1) = (1, 0)$$

Therefore,  $\boxed{\pi(e_1) = (1, 0)}$  point in the direction of the  $x$ -axis.

19. a

Let  $w_1, w_2$  be the orthonormal basis of  $W$  with respect to which we have

$$\pi(e_i) = (a_i, b_i)^t, \quad i = 1, 2, 3,$$

where  $e_i$  are the standard basis vectors of  $\mathbb{R}^3$ . First let us extended  $w_1, w_2$  to an orthonormal basis of  $\mathbb{R}^3$ , say  $w_1, w_2, w_3$ . Then if we write a vector  $x \in \mathbb{R}^3$  in the basis composed of  $w_i$ , i.e.  $x = (x_1, x_2, x_3) = x_1 w_1 + x_2 w_2 + x_3 w_3$ , then our orthogonal projection has a simple form,

$$\pi(x) = \pi((x_1, x_2, x_3)) = (x_1, x_2). \quad (1)$$

Thus we want to write vectors  $e_1, e_2, e_3$  in the basis consisting of  $w_i$ s. In order to do this, note that if

$$W = [w_1 \quad w_2 \quad w_3]$$

, where  $w_i$  are thought of as column vectors (and hence  $W$  is a  $3 \times 3$  matrix), then

$$W e_i = w_i.$$

Therefore,

$$e_i = W^{-1} w_i \quad (2)$$

is the expression of  $e_i$  in the basis  $w_1, w_2, w_3$ . Furthermore we see from (2) that coordinate vector of  $e_i$  in this basis is the  $i$ th column of  $W^{-1}$ . Now by (1) we see that  $(a_1, a_2, a_3)$  is the first row of  $W^{-1}$  and  $(b_1, b_2, b_3)$  is the second row of  $W^{-1}$ . But note that since  $W$  is orthogonal then  $W^{-1}$  also is, since  $W^{-1} = W^t$ ; hence its columns and rows are orthogonal unit vectors, showing that  $(a_1, a_2, a_3)$  and  $(b_1, b_2, b_3)$  are indeed orthogonal unit vectors.

## Result

2 of 2

First we extended the orthonormal basis of  $W$  to orthonormal basis of  $\mathbb{R}^3$  and show that the projection formula has a simple form in case when we write  $e_i$  in that basis, and then we use that coupled with orthogonality of the basis of  $W$  to show our desired result.

20. a

**Theorem 8.2.5:**

The following properties of a real  $n \times n$  matrix  $A$  are equivalent:

- (i) The form  $X^T A Y$  represents dot product, with respect to some basis of  $\mathbb{R}^n$ .
- (ii) There is an invertible matrix  $P$  such that  $A = P^T P$ .
- (iii) The matrix  $A$  is symmetric and positive definite.

**Prove that** the form and the matrix are positive definite if and only if  $\det A_k > 0$  for  $k = 1, \dots, n$ .

Suppose that matrix  $A$  is positive definite.

If  $A$  is  $1 \times 1$ , then  $u^T A u = a_{11} u^2 > 0$  if and only if  $a_{11} > 0$  and  $v \neq 0$ .

So, the criterion holds for  $n = 1$ .

Let  $n \geq 2$

Assume the criterion holds for  $1, \dots, n-1$ , and let  $A$  be a  $n \times n$  symmetric real matrix.

If  $A$  is positive definite, the matrix  $A$  reduce to  $A = P^T P$  for some invertible matrix  $P$  by theorem 8.2.5.

Then,

$$\det A = (\det P)^2 > 0$$

Since  $A_{n-1}$  is positive definite. Apply the inductive hypothesis to deduce that  $\det A_k > 0$  for  $k = 1, \dots, n-1$ .

Therefore,  $\det A_k > 0$  for  $k = 1, \dots, n$

**Conversely,**

Suppose that  $\det A_k > 0$  for  $k = 1, \dots, n-1$ . By induction,  $A_{n-1}$  defines a positive definite form on the subspace  $S$ .

Therefore,  $A_{n-1} = Q^T Q$  for some invertible  $(n-1) \times (n-1)$  matrix.

Another way of saying this is that there exists an orthonormal basis  $\{s_1, \dots, s_{n-1}\}$  of  $S$  with change of basis matrix  $M$ .

So,

$$A_{n-1} = M^T I M$$

Extend this basis to an orthogonal basis of  $\mathbb{R}^n$ :

Let  $v$  be in the complement of  $S$  and let  $u = v - \pi(u)$ , where  $\pi$  the orthogonal projection to is  $S$ . Because the form defined by  $A$  is non-degenerate, so a projection exists.

Since,

$$\det A \neq 0$$

In this way, to obtain an orthogonal basis  $\{s_1, \dots, s_{n-1}, u\}$

Thus, there is a change of basis matrix  $P$  such that  $A = P^T D P$ , where  $D$  is diagonal and has  $(n-1)$  ones on the diagonal and  $d_{nn} \in \{-1, 0, 1\}$ .

Since,

$$\det D = (\det A)(\det P)^2 > 0,$$

Then,

$$d_{nn} = 1$$

Therefore,  $A = P^T P$ , and so,  $A$  is positive definite.

**Hence proved**

21. a

**Statement of Sylvester's law:**

The signature of symmetric form on a real vector space or of a Hermitian form on a complex vector space does not depend on the choice of orthogonal basis.

[Comment](#)

Step 2 of 5 ^

**Proof of Sylvester's law:**

Suppose that  $\mathbf{B} = (v_1, \dots, v_d)$  and  $\mathbf{B}' = (v'_1, \dots, v'_d)$  are two ordered orthogonal bases for a real or complex vector space  $V$  which is provided with a symmetric or Hermitian form  $\langle \cdot, \cdot \rangle$ .

Suppose that the bases are ordered, so that the matrices  $M$  and  $M'$  of the form have positive entries, followed by negative entries and zero entries along the diagonal.

Let  $(q, l, n)$  and  $(q', l', n')$  be the signatures of these matrices.

Since these matrices are in reduced row echelon form, it can be seen that the ranks are equal to  $d - n$  and  $d - n'$ .

So, the ranks of  $M$  and  $M'$  are equal.

Since there is  $d \times d$  matrix  $S$  which is invertible such that  $S^*MS = M'$ .

Therefore,  $n = n'$

To show that  $q = q'$  and, hence that  $l = l'$ .

Suppose to the contrary that  $q > q'$ .

Then,

$$l < l'$$

Let subspace  $Q$  be spanned by  $(v_1, \dots, v_q)$ , so that,  $\langle \cdot, \cdot \rangle$  is positive definite on  $Q$ .

Let subspace  $L'$  be the spanned by  $(v'_{q'+1}, \dots, v'_{l'})$ , so that  $\langle \cdot, \cdot \rangle$  is negative definite on  $L'$ .

Let two subspaces  $N$  and  $N'$  be the spanned by the last  $n = n'$  vectors of the bases  $\mathbf{B}$  and  $\mathbf{B}'$ .

If there were a vector  $v$  in  $N \cap N'$ , then the span of  $N' \cup \{v\}$  would be a vector space on which the form was zero, contradicting the fact because the dimension of such a subspace is at most  $d - n = d - n'$ .

Therefore,  $N = N'$

Consider  $W = (Q + N) \cap (L' + N')$ .

This is a vector space of dimension at least  $n + 1$ .

If  $w \in W - N$ , then  $w \in Q \cap L'$ .

Since  $\langle w, w \rangle$  is both positive definite and negative definite on  $Q \cap L'$ .

It must have  $w = 0$ .

This implies that, the dimension of  $W$  is  $n$ , which is a contradiction.

Therefore,  $q = q'$  and  $n = n'$ .

**Hence**, the signature of symmetric form on a real vector space or of a Hermitian form on a complex vector space does not depend on the choice of orthogonal basis.

## Section 5



1. a

(a)

Note that since  $V$  is a Euclidean space, we can choose a basis of  $V$  with respect to which  $\langle \cdot, \cdot \rangle$  is a dot product. This is a consequence of **Theorem 8.2.5**, which says that it is a sufficient and necessary condition for a form  $X^t A Y$  to represent dot product, is that the matrix  $A$  is positive definite symmetric, which is given by the fact that  $V$  is a Euclidean space. Thus choose a basis with respect to which our bilinear form is a dot product, and let

$$X = [x_1 \quad \cdots \quad x_n]^t$$

and

$$Y = [y_1 \quad \cdots \quad y_n]^t$$

be column vectors with respect of that basis of  $v$  and  $w$ , respectively. Then we have to prove that

$$|X^t Y| \leq \sqrt{X^t X} \sqrt{Y^t Y}. \quad (1)$$

As both sides are positive and squaring is a monotone increasing function on positive numbers, then (1) is equivalent to

$$(X^t Y)^2 \leq (X^t X)(Y^t Y).$$

Writing it out we see that that is equivalent to

$$\left( \sum_{i=1}^n x_i y_i \right)^2 \leq \left( \sum_{i=1}^n x_i^2 \right) \left( \sum_{i=1}^n y_i^2 \right). \quad (2)$$

Now we consider the polynomial

$$p(t) = \sum_{i=1}^n (x_i t + y_i)^2 = \left( \sum_{i=1}^n x_i^2 \right) t^2 + 2 \left( \sum_{i=1}^n x_i y_i \right) t + \left( \sum_{i=1}^n y_i^2 \right).$$

As  $p(t)$  is a sum of squares then

$$p(t) \geq 0$$

for all real  $t$ ; but this means that  $p(t)$  has at most one real root. Recall that a real quadratic polynomial has 0 or 1 ("double") real root if and only if its discriminant is less or equal to 0. But noting that the discriminant of a polynomial  $P(x) = ax^2 + bx + c$  is  $b^2 - 4ac$  then we see that the discriminant of  $p(t)$  is

$$\left( 2 \left( \sum_{i=1}^n x_i y_i \right) \right)^2 - 4 \left( \sum_{i=1}^n x_i^2 \right) \left( \sum_{i=1}^n y_i^2 \right), \quad (3)$$

where by noting that (3) is less than or equal to 0, rearranging and dividing through with 4 we get precisely (2).

(b)

By the definition of  $|\cdot|$  we have that

$$|v + w|^2 = \langle v + w, v + w \rangle \text{ and } |v - w|^2 = \langle v - w, v - w \rangle.$$

Now by using this and bilinearity and symmetry of the form we have

$$\begin{aligned} |v + w|^2 + |v - w|^2 &= \langle v + w, v + w \rangle + \langle v - w, v - w \rangle \\ &= \langle v, v + w \rangle + \langle w, v + w \rangle + \langle v, v - w \rangle - \langle w, v - w \rangle \\ &= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle + \langle v, v \rangle - \langle v, w \rangle - \langle w, v \rangle + \langle w, w \rangle \\ &= 2\langle v, v \rangle + 2\langle w, w \rangle = 2|v|^2 + 2|w|^2, \end{aligned}$$

which is what we need to prove.

(c)

Recall that  $|v| = |w|$  means that  $\sqrt{\langle v, v \rangle} = \sqrt{\langle w, w \rangle}$ . But as  $\langle v, v \rangle$  and  $\langle w, w \rangle$  are both nonnegative real numbers (as the form is positive definite), then this implies that

$$\langle v, v \rangle = \langle w, w \rangle. \quad (3)$$

Now we compute

$$\begin{aligned} \langle v + w, v - w \rangle &= \langle v, v - w \rangle + \langle w, v - w \rangle \\ &= \langle v, v \rangle - \langle v, w \rangle + \langle w, v \rangle - \langle w, w \rangle \\ &= 0 \end{aligned}$$

where in the last equality we used the symmetry of the form and (1).

## Result

4 of 4

In (a) part we show that it is sufficient to show it for dot product, and then show it directly. In (b) part we write out everything according to the definition of the length of a vector, and then the claim follows through some simple algebra; same happens in (c) after we note that  $|v| = |w|$  implies  $\langle v, v \rangle = \langle w, w \rangle$ .

2. a

We have

$$W^\perp = \{v \in V : \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

Then

$$\begin{aligned} W^{\perp\perp} &= \{v \in V : \langle v, w' \rangle = 0 \text{ for all } w' \in W^\perp\} \\ &= \{v \in V : \text{there exists } w' \in W^\perp \text{ such that } \langle v, w' \rangle = 0 \text{ and } \langle w, w' \rangle = 0 \text{ for all } w \in W\}. \end{aligned}$$

First note immediately that any element of  $W$  satisfies this definition, for any  $w' \in W^\perp$  fits the definition and both equalities are satisfied for  $v \in W$ , hence  $W \subseteq W^{\perp\perp}$ .

Now note by **Corollary 8.5.1** we have that that

$$V = W \oplus W^\perp$$

and

$$V = W^\perp \oplus W^{\perp\perp}.$$

(Since  $W^\perp$  is a subspace of  $V$  and thus nondegenerate.)

Thus the dimensions of  $W$  and  $W^{\perp\perp}$  are the same, and hence equality holds.

## Result

2 of 2

First we show that  $W \subset W^{\perp\perp}$ , and then we use the fact that the form on Euclidean space is nondegenerate on any subspace in order to obtain that  $W$  and  $W^{\perp\perp}$  have the same dimensions, and thus they're equal.

3. a

Let  $w \in \mathbb{R}^n$  be a vector of length 1, and let  $U$  denote the orthogonal space  $w^\perp$ .

The reflection  $r_w$  about  $U$  is defined as follows: write a vector  $v$  in the form  $v = cw + u$

where  $u \in U$ . Then  $r_w(v) = -cw + u$

[Comment](#)

Step 2 of 5 ^

(a)

Prove that the matrix  $P = I - 2ww^T$  is orthogonal.

First show that  $P^T = P$ ;

$$\begin{aligned} P^T &= (I - 2ww^T)^T \\ &= I^T - (2ww^T)^T \\ &= I - 2(w^T)^T w \\ &= I - 2ww^T \\ &= P \end{aligned}$$

Now,

$$\begin{aligned} P^T P &= I - 4ww^T + 4ww^T \\ &= I \end{aligned}$$

Therefore, the matrix  $P = I - 2ww^T$  is orthogonal.

**Hence proved**

(b)

Prove that multiplication by  $P$  is a reflection about the orthogonal space  $U$ .

Let a vector  $v \in \mathbb{R}^n$  can be written as  $v = cw + u$  where  $u \in U$

Then,

$$\begin{aligned} w^T v &= w^T (cw + u) \\ &= w^T cw + w^T u \\ &= cw^T w + w^T u \end{aligned}$$

Since,  $w \in \mathbb{R}^n$  be a vector of length 1 and  $U$  denote the orthogonal space  $w^\perp$ .

Then,  $w^T w = 1$  and  $w^T u = 0$

$$\begin{aligned} w^T v &= c \cdot 1 + 0 \\ &= c \end{aligned}$$

Now computing  $Pv$ :

$$\begin{aligned} Pv &= (I - 2ww^T)v \\ &= I \cdot v - 2ww^T v \\ &= v - 2wc \end{aligned}$$

Substitute  $v = cw + u$  into the above, to obtain

$$\begin{aligned} Pv &= cw + u - 2cw \\ &= -cw + u \\ &= r_w(v) \end{aligned}$$

Hence proved

[Comment](#)

Step 5 of 5 ^

(c)

Let  $u, v$  be vectors of equal length in  $\mathbb{R}^n$ . Determine a vector  $w$  such that  $Pu = v$ .

Since  $u, v$  are vectors of equal length in  $\mathbb{R}^n$  then assume that a vector  $w = u - v$ , normalized to be length 1 such that  $Pu = v$ .

Therefore, a vector  $w = \frac{u - v}{\|u - v\|}$

4. a

(a)

We first prove that  $\ker T$  and  $\text{im } T$  are orthogonal subspaces of  $V$ . We must prove that for each  $v \in \ker T$  and  $w \in \text{im } T$  we have that  $\langle v, w \rangle = 0$ . In particular, let  $X$  and  $Y$  be coordinate column vectors of  $v$  and  $w$  according to the standard basis of  $\mathbb{R}^n$ , then we have that  $AX = 0$  and  $AY' = Y$  for some  $Y'$ . Hence we compute

$$\begin{aligned} \langle v, w \rangle &= X^t Y \\ &= X^t (AY') \\ &= (X^t A) Y' \\ &= (A^t X)^t Y' \\ &= (AX)^t Y' \\ &= 0 \end{aligned}$$

where the second to last equality follows from the fact that since  $A$  is symmetric then  $A^t = A$ , and the last line follows from  $AX = 0$ . Thus we have that  $\ker T$  and  $\text{im } T$  are orthogonal. Recall by rank-nullity theorem that we have

$$\dim(\ker T) + \dim(\text{im } T) = \dim(V),$$

and hence in order to conclude that  $V$  is a direct sum it is sufficient to show that the intersection of  $\ker T$  and  $\text{im } T$  is  $\{0\}$ . Now let  $v \in \ker T$  and  $v \in \text{im } T$ , then since  $\ker T$  and  $\text{im } T$  are orthogonal, we have  $\langle v, v \rangle = 0$ ; but as we are in Euclidean space this implies that  $v = 0$ .

(b)

Suppose first that  $T$  is an orthogonal projection onto  $\text{im } T$ . Recall that an orthogonal projection onto  $\text{im } T$  is a mapping  $\pi : V \rightarrow \text{im } T$  such that if  $v \in \text{im } T$  and  $w \in \ker T$ , then

$$\pi(v + w) = v.$$

Hence we have

$$T(v + w) = v,$$

but note that since  $T$  is linear mapping then  $T(v + w) = T(v) + T(w) = T(v)$ , and hence this means that we have

$$T(v) = v.$$

This implies that

$$T \circ T(v) = T(v),$$

where  $A^2$  is the matrix of the mapping on the left-hand side and  $A$  is the matrix of the mapping on the right-hand side. Note also that for  $w \in \ker T$ ,  $T \circ T(w) = T(T(w)) = T(0) = 0$ , and hence we must have that  $T \circ T$  and  $T$  are the same linear mapping; but this is possible only if

$$A^2 = A.$$

This proves the " $\Rightarrow$ " part of the implication. Now suppose  $A^2 = A$ . Then, as  $A^2$  is the matrix of the linear mapping  $T \circ T$  and  $A$  is the matrix of the mapping  $T$ , then for any  $k \in V$  we have

$$T \circ T(k) = T(T(k)) = T(k). \quad (1)$$

If  $k$  is in  $\ker T$  then obviously (1) equals 0. Now if  $p \in \text{im } T$ , then there exists an  $t \in T$  such that  $T(t) = p$ , and hence (1) implies

$$T(T(t)) = T(t) \text{ and hence } T(p) = p.$$

But this suffices to demonstrate that if  $k \in \ker T$  and  $p \in \text{im } T$ , then  $T(p + k) = p$ , which is what we needed to show in order to prove that  $T$  is an orthogonal transformation.

## Result

3 of 3

In the (a) part we use the definition of  $\ker T$  and  $\text{im } T$  and the fact that since  $A$  is symmetric then  $A^t = A$ , in order to show that if  $v \in \ker T$  and  $w \in \text{im } T$  then  $\langle v, w \rangle = 0$ .

For the (b) part we use the fact that  $A^2$  is the matrix associated with the linear mapping  $T \circ T$  in order to obtain our desired results. Click for more details.

5. a



Recall that  $P$  being unitary means that  $P^*P = I$ . We also have that

$$PX_1 = \lambda_1 X_1 \text{ and } PX_2 = \lambda_2 X_2.$$

Let us first show that

$$|\lambda_1|^2 = \lambda_1 \overline{\lambda_1} = 1 \quad (1)$$

(and therefore  $|\lambda_1| = \sqrt{1} = 1$ ; also by symmetry the same holds for  $\lambda_2$ ; this is the fact that all the eigenvalues of a unitary matrix are on the complex unit circle).

We have

$$\begin{aligned} X_1^* X_1 &= X_1^* (P^* P) X_1 \\ &= (X_1^* P^*) (P X_1) \\ &= (P X_1)^* (P X_1) \\ &= (\lambda_1 X_1)^* (\lambda_1 X_1) \\ &= \overline{\lambda_1} \lambda_1 (X_1^* X_1), \end{aligned}$$

showing (1) as eigenvectors are nonzero and the standard Hermitian form is positive definite, and hence  $X_1^* X_1 = \langle X_1, X_1 \rangle > 0$ .

Now we compute

$$\begin{aligned} \langle X_1, X_2 \rangle &= X_1^* X_2 \\ &= X_1^* (P^* P) X_2 \\ &= (P X_1)^* (P X_2) \\ &= (\lambda_1 X_1)^* (\lambda_2 X_2) \\ &= \overline{\lambda_1} \lambda_2 (X_1^* X_2), \end{aligned}$$

showing that either  $\langle X_1, X_2 \rangle = 0$ , in which case we are done, or

$$\overline{\lambda_1} \lambda_2 = 1. \quad (2)$$

But noting that we have proved (1), and multiplying both sides of (2) by  $\lambda_1$ , we obtain

$$\lambda_2 = \lambda_1,$$

which is contrary to  $\lambda_1$  and  $\lambda_2$  being distinct eigenvalues, and hence  $\langle X_1, X_2 \rangle = 0$ , i.e.  $X_1$  and  $X_2$  are orthogonal.

## Result

2 of 2

First we prove that all the eigenvalues of a unitary matrix are complex numbers  $\lambda$  with  $|\lambda| = 1$ , and then we use this to compute  $\langle X_1, X_2 \rangle$  as zero. Click for more details.

6. a



Let  $P$  be a unitary matrix and  $\lambda$  its eigenvalue, i.e.  $PX = \lambda X$  for some column vector  $X$ .

Then we have

$$(\lambda X)^*(\lambda X) = \bar{\lambda} X^* \lambda X = (\bar{\lambda} \lambda) X^* X. \quad (1)$$

But we also have

$$(\lambda X)^*(\lambda X) = (PX)^* PX = X^* P^* PX = X^* X. \quad (2)$$

Now since  $X \neq 0$ , it follows immediately from (1) and (2) that

$$\bar{\lambda} \lambda = 1. \quad (3)$$

Recall that for a complex number  $z$  we have  $|z| = \bar{z}z$ , and hence (3) can be stated as requirement that  $\lambda$  has modulus 1.

## Result

2 of 2

We show that any eigenvalue of a unitary matrix is a complex number with modulus 1 (i.e. it lies in the unit circle on the complex plane).

# Section 6

1. a

## Proof of (c) part of the proposition

Suppose  $T$  is Hermitian, then

$$T = T^*. \quad (1)$$

Note that it was prove in (a) that for all  $v$  and  $w$  in  $V$  we have

$$\langle Tv, w \rangle = \langle v, T^*w \rangle,$$

but using (1) here this transforms into

$$\langle Tv, w \rangle = \langle v, Tw \rangle,$$

which is what we wanted to prove.

Conversely, assume that we have, for all  $v$  and  $w$  in  $V$ ,

$$\langle Tv, w \rangle = \langle v, Tw \rangle.$$

Again by using (a), namely that  $\langle v, Tw \rangle = \langle T^*v, w \rangle$ , we get

$$\langle Tv, w \rangle = \langle T^*v, w \rangle$$

and hence we have that

$$\langle Tv - T^*v, w \rangle = 0$$

holds for all  $v$  and  $w$ . Then as the form on Hermitian space is nondegenerate we can apply **Proposition 8.4.3** to show that  $Tv - T^*v = 0$  for all  $v \in V$ , or equivalently that

$$Tv = T^*v$$

for all  $v$ , which is what we wanted to prove.

### Proof of (d) part of the proposition

Suppose  $T$  is unitary, then

$$TT^* = T^*T = I. \quad (2)$$

Now we write

$$\langle Tv, Tw \rangle = \langle v, T^*Tw \rangle = \langle v, Iw \rangle = \langle v, w \rangle$$

where in the first equality we used the fact that  $\langle Tv, w \rangle = \langle v, T^*w \rangle$ , which was proven in (a).

Let us now prove the converse, so suppose that we have

$$\langle Tv, Tw \rangle = \langle v, w \rangle$$

for all  $v$  and  $w$  in  $V$ . Then by (a) we have that

$$\langle Tv, Tw \rangle = \langle T^*Tv, w \rangle$$

and hence

$$\langle T^*Tv, w \rangle = \langle v, w \rangle$$

which implies that for

$$\langle T^*Tv - v, w \rangle = 0.$$

By the nondegeneracy of the form on Hermitian space, analogously as in the (c) part, this implies that

$$T^*Tv = v,$$

i.e.  $T^*T = I$  and hence  $T$  is unitary.

### Result

3 of 3

We prove both (c) and (d) parts by applications of (a) and by nondegeneracy of the form on Hermitian space  $V$  implying that certain suitably chosen vectors are equal. Click to see more details.

### 2. a

As  $T$  is a symmetric operator then according to the cited proposition we have that for any  $w$  and  $w'$ ,

$$\langle T(w), w' \rangle = \langle w, T(w') \rangle.$$

Putting  $w' = T(v)$  we get

$$\langle T(w), T(v) \rangle = \langle w, T(T(v)) \rangle = \langle w, 0 \rangle = 0.$$

As this holds for any  $w$ , then by setting  $w = v$  we obtain

$$\langle T(v), T(v) \rangle = 0,$$

and as we are in a Euclidean space (where the form is positive definite), this implies that  $T(v) = 0$ .

### Result

2 of 2

We use the cited proposition to show that  $\langle T(v), T(v) \rangle = 0$ , and hence this implies  $T(v) = 0$ . Click for more details.

### 3. a

Let  $A$  be a symmetric and orthogonal real  $3 \times 3$  matrix. As  $A$  is symmetric then we have  $A = A^t$ , and hence, by Spectral Theorem for symmetric matrices, there exists an orthogonal matrix  $P$  such that

$$P^t A P = P^t A^t P = D,$$

and  $D$  is a diagonal real matrix. Therefore we also have

$$\begin{aligned} (P^t A P)(P^t A^t P) &= (P^t A (P P^t) A^t P), \\ &= (P^t A A^t P) \\ &= P^t P \\ &= I \end{aligned}$$

where in the second and fourth equalities we used the fact that  $P$  is orthogonal, and in the third one we used the fact that  $A$  is orthogonal. Hence  $D^2 = I$ , and as  $A$  and  $D$  are similar matrices and thus have the same eigenvalues, this implies that all of the eigenvalues of  $A$  are equal to  $\pm 1$ .

Hence a real symmetric and orthogonal  $3 \times 3$  matrix is similar to one (and only one) of these matrices:

$$I, -I, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

#### Result

2 of 2

It tells us that it is diagonalizable as a diagonal matrix with all entries equal to  $\pm 1$ , and thus that all its eigenvalues are equal to  $1$  or  $-1$ . We also list an exhaustive list of the diagonal matrices with entries equal to  $\pm 1$  to which it can be similar. Click to see more details.

### 4. a

Let  $A$  be an  $n \times m$  matrix such that  $A^* A = D$  is diagonal. If we write

$$A = [a_1 \ a_2 \ \cdots \ a_m]$$

where  $a_i$  is the  $i$ th column of  $A$ . Then we have that

$$A^* A = \begin{bmatrix} \overline{a_1}^t \\ \overline{a_2}^t \\ \vdots \\ \overline{a_m}^t \end{bmatrix} [a_1 \ a_2 \ \cdots \ a_m] = \begin{bmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle & \cdots & \langle a_1, a_m \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle & \cdots & \langle a_2, a_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle a_m, a_1 \rangle & \langle a_m, a_2 \rangle & \cdots & \langle a_m, a_m \rangle \end{bmatrix}$$

where the form is the standard Hermitian form. Therefore since  $A^* A$  is diagonal it follows that  $\langle a_i, a_j \rangle = 0$ , for all  $i \neq j$ , and hence the columns of  $A$  are orthogonal.

#### Result

2 of 2

We show that the columns of  $A$  are orthogonal. Click for more details.

### 5. a

$A = (a_{ij})$  is a real skew-symmetric matrix if and only if it is a real  $n \times n$  matrix such that  $a_{ij} = -a_{ji}$  for all  $1 \leq i, j \leq n$ , i.e.  $A^t = -A$ . Then we have that  $iA = (ia_{ij})$ , and hence by the definition of adjoint matrix we have

$$(iA)^* = \bar{i}A^* = -iA^t = -i(-A) = iA,$$

which is what we needed prove. Now we can apply the Spectral theorem for Hermitian matrices to  $iA$ , which states that there exists an orthogonal matrix  $P$  such that

$$P^*(iA)P = D$$

is a real diagonal matrix, but we have

$$P^*AP = -iD$$

(as the multiplicative inverse of  $i$  is  $-i$ , and hence  $A$  is similar to a diagonal matrix with entries being either zero or purely imaginary (i.e. nonzero complex numbers with real part 0)). As similar matrices have the same eigenvalues, this also implies that eigenvalues of  $A$  are purely imaginary.

## Result

2 of 2

We prove the desired result through the use of basic properties of skew-symmetric matrices and matrix adjoints, and then apply the Spectral theorem in order to show that a skew-symmetric matrix are complex numbers with real part 0, i.e. purely imaginary.

## 6. a

Suppose that  $A$  is invertible and normal, i.e.

$$A^*A = AA^*. \quad (1)$$

First note that we have  $(A^{-1})^* = (A^*)^{-1}$  -- this is an easy consequence of taking an adjoint of both sides in the equality  $A^{-1}A = I$ . Then we also see that  $A$  is invertible if and only if  $A^*$  is invertible. Now, let us transform (1) by taking an inverse of both sides; then this turns into

$$A^{-1}(A^*)^{-1} = (A^*)^{-1}A^{-1}. \quad (2)$$

Recall now that  $Q$  is unitary if and only if  $QQ^* = I$ . Now we compute

$$\begin{aligned} (A^*A^{-1})(A^*A^{-1})^* &= A^*A^{-1}(A^{-1})^*A \\ &= A^*A^{-1}(A^*)^{-1}A \\ &= A^*(A^*)^{-1}A^{-1}A \\ &= I, \end{aligned}$$

where we used (2) (i.e. normality) in third equality.

Conversely, suppose that  $A^*A^{-1}$  is unitary, then

$$A^*A^{-1}(A^{-1})^*A = I.$$

Via multiplying by  $A^{-1}$  from the right and  $(A^*)^{-1}$  from the left this turns into

$$A^{-1}(A^{-1})^* = (A^*)^{-1}A^{-1},$$

which is, after correcting for the order in which we take adjoints and inverses, equivalent with (2), and thus we obtain normality of  $A$  by taking an inverse of both sides.

## Result

3 of 3

After noting that invertibility of  $A$  implies the invertibility of  $A^*$  and that  $(A^{-1})^* = (A^*)^{-1}$ , the result follows from some algebraic manipulations of matrix identities. Click to see more details.

## 7. a

Since  $P$  is normal, by the Spectral Theorem for normal matrices we have that there exists a unitary matrix  $Q$  such that

$$Q^*PQ = D,$$

where  $D$  is a diagonal matrix. Note that since  $P$  and  $D$  are similar then they have the same eigenvalues, and hence  $D$  is a real matrix. Now we have (since  $Q^* = Q^{-1}$ ) that

$$P = QDQ^*.$$

Taking an adjoint of both sides this implies that

$$P^* = QD^*Q^*,$$

but as  $D$  is a real diagonal matrix then  $D = D^*$  and consequently  $P = P^*$ . Since  $P$  is real then  $P^* = P^t$ , and this completes our proof, as  $P$  is a symmetric matrix if and only if  $P = P^t$ .

## Result

2 of 2

We use the Spectral theorem for normal matrices and characterization of eigenvalues of a diagonal matrix in order to prove our result. Click for more details.

## 8. a

### (a)

Conjugate linearity in the first variable and linearity in the second variable both follows from the linearity of the integral, while the fact that it is Hermitian symmetric follows from noting that

$$\overline{\int_0^{2\pi} f(\theta)g(\theta)d\theta} = \int_0^{2\pi} \overline{f(\theta)g(\theta)}d\theta = \int_0^{2\pi} \overline{g(\theta)}\overline{f(\theta)}d\theta.$$

It remains to show that it is positive definite, but note that

$$\int_0^{2\pi} \overline{f(\theta)}f(\theta)d\theta = \int_0^{2\pi} |f(\theta)|^2d\theta \geq 0 \quad (1)$$

because  $|f(\theta)|^2 \geq 0$ . As  $f$  is differentiable (and hence also continuous) then if it is nonzero on some point  $x \in [0, 2\pi]$  then it is also nonzero on some  $\epsilon$  neighbourhood of  $x$ , and therefore (1) is strictly positive.



(b)

Note that for any nonzero integer  $k$  we have

$$\int_0^{2\pi} e^{k\theta i} d\theta = 0, \quad (2)$$

and also note that  $\overline{e^{i\theta}} = e^{-i\theta}$  -- these follow from the Euler's formula  $e^{i\theta} = \cos \theta + i \sin \theta$ .

It is not difficult to see that  $W$  has dimension  $n + 1$ , since a scalar combination of terms with  $e^{i\theta k}$  with  $k < m$  cannot equal  $e^{i\theta m}$ . Furthermore, let

$$f(x) = \sum_{k=0}^n a_k x^k \quad \text{and} \quad g(x) = \sum_{k=0}^n b_k x^k$$

be polynomials. Then we have

$$\begin{aligned} \int_0^{2\pi} \overline{f(e^{i\theta})} g(e^{i\theta}) d\theta &= \int_0^{2\pi} \overline{\sum_{k=0}^n a_k (e^{i\theta})^k} \left( \sum_{k=0}^n b_k (e^{i\theta})^k \right) d\theta \\ &= \int_0^{2\pi} \left( \sum_{k=0}^n \overline{a_k (e^{i\theta})^k} \right) \left( \sum_{k=0}^n b_k (e^{i\theta})^k \right) d\theta \\ &= \int_0^{2\pi} \left( \sum_{k=0}^n \overline{a_k} e^{-i\theta k} \right) \left( \sum_{k=0}^n b_k e^{i\theta k} \right) d\theta \\ &= \int_0^{2\pi} \sum_{k=0}^n \overline{a_k} b_k d\theta \\ &= \sum_{k=0}^n \overline{a_k} b_k \int_0^{2\pi} d\theta \\ &= 2\pi \sum_{k=0}^n \overline{a_k} b_k, \end{aligned} \quad (3)$$

where the fourth equality is application of (2). As it seems that real scalars would suffice, let  $a_k$  be real, then  $\overline{a_k} = a_k$ ; for the normality we then require that  $\sum_{k=0}^n a_k^2 = \frac{1}{2\pi}$ , while orthogonality between  $f$  and  $g$  means that  $\sum_{k=0}^n a_k b_k = 0$ . It is now not difficult to verify that the polynomials

$$\frac{1}{\sqrt{2\pi}}, \frac{x}{\sqrt{2\pi}}, \dots, \frac{x^n}{\sqrt{2\pi}}$$

satisfy these requirements and that they're independent (by the considerations after (2)).



(c)

We want to prove that for any  $f$  and  $g$  in  $V$  we have

$$\langle T(f), g \rangle = \langle f, T(g) \rangle,$$

as this is equivalent to  $T$  being Hermitian by **Proposition 8.6.3**. Hence we compute

$$\begin{aligned}\langle T(f), g \rangle &= \int_0^{2\pi} i \overline{f'(\theta)} g(\theta) d\theta \\ &= \int_0^{2\pi} -i \overline{f'(\theta)} g(\theta) d\theta \\ &= -i \int_0^{2\pi} \overline{f'(\theta)} g(\theta) d\theta\end{aligned}$$

where note that the integral is just integration by parts, since  $\overline{f'(\theta)} = \overline{f'}(\theta)$ . Hence we have

$$\begin{aligned}\langle T(f), g \rangle &= -i \int_0^{2\pi} \overline{f'(\theta)} g(\theta) d\theta \\ &= -i \left[ \overline{f(\theta)} g(\theta) \right]_0^{2\pi} + i \int_0^{2\pi} \overline{f(\theta)} g'(\theta) d\theta\end{aligned}$$

where note that  $\left[ \overline{f(\theta)} g(\theta) \right]_0^{2\pi} = 0$  because, as we're on the circle,  $f$  and  $g$  have period  $2\pi$ . Therefore,

$$\begin{aligned}\langle T(f), g \rangle &= i \int_0^{2\pi} \overline{f(\theta)} g'(\theta) d\theta \\ &= \int_0^{2\pi} \overline{f(\theta)} i g'(\theta) d\theta \\ &= \langle f, T(g) \rangle,\end{aligned}$$

which is what we needed to show.

In order to find the eigenvalues, we need to find all  $\lambda$  such that there exists  $f \in W$  so that

$$if'(\theta) = \lambda f(\theta).$$

This is just a simple first-order ordinary differential equation, which after writing as

$$\frac{f'(\theta)}{f(\theta)} = -i\lambda$$

we can integrate in order to obtain

$$f(\theta) = ce^{-i\lambda\theta},$$

for some positive constant  $c$ . As we're on the circle we require  $f(\theta) = f(\theta + 2\pi)$ , and hence

$$ce^{-i\lambda\theta} = ce^{-i\lambda\theta - i\lambda 2\pi}$$

which implies

$$ce^{-i\lambda 2\pi} = 1,$$

which can be seen (by e.g. Euler's formula) to imply that  $\lambda \in \mathbb{Z}$ .

We also require  $f$  to be an element of  $W$ , and hence  $f(\theta)$  is of the form  $\sum_{k=0}^n a_k e^{ik\theta}$ , so that  $e^{-i\lambda\theta}$ , for an integer  $\lambda$ , is an element of  $W$  if and only if  $\lambda = 0, -1, \dots, -n$ .

## Result

5 of 5

Most properties in **(a)** follow from linearity of conjugation and linearity of the integral, as well as differentiability of  $f \in V$ . Next in **(b)** we find that polynomials  $\frac{1}{\sqrt{2\pi}}, \frac{x}{\sqrt{2\pi}}, \dots, \frac{x^n}{\sqrt{2\pi}}$  form an orthonormal basis, and in **(c)** that  $0, -1, \dots, n$  are eigenvalues on  $W$ .

9. a

Note that

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is symmetric, and hence we use the fact that we can diagonalize  $A$  by multiplying it with a matrix whose columns are eigenvectors of  $A$ , which is consequence of the Spectral theorem for symmetric operators. Therefore, first we find an eigenbasis of  $\mathbb{R}^2$ , which is not hard by noting that

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} \quad (1)$$

Thus if

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

then  $AX_1 = X_1$ , while if

$$X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

then  $AX_2 = -X_2$ , showing that 1 and  $-1$  are eigenvalues of  $A$  corresponding to eigenvectors  $X_1$  and  $X_2$ , respectively -- it is also immediate that  $X_1$  and  $X_2$  form a basis of  $\mathbb{R}^2$ . We see that both  $X_1$  and  $X_2$  have the (usual Euclidean) length 2, and thus we can normalize them by dividing them by  $\sqrt{2}$ . Furthermore the matrix

$$P = [X_1/\sqrt{2} \quad X_2/\sqrt{2}]$$

is orthogonal and we have

$$P^t A P = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Now, note that  $X_1$  and  $X_2$  are orthogonal according to the form induced by  $A$ , and recall the matrix form of the corollary giving the existence of the signature and the Sylvester's law stating that the signature does not depend on the choice of orthogonal basis -- therefore the signature of the form is  $(1, 1)$ .

## Result

2 of 2

We use the methods outlined in this chapter in order to find such an orthogonal matrix, and in the process we find that the signature of the form is  $(1, 1)$ .

10. a

In order to check that  $\{, \}$  is a Hermitian form, we first check that it is additive in both coordinates, i.e.

$$\{v_1 + v_2, w\} = \{v_1, w\} + \{v_2, w\} \text{ and } \{v, w_1 + w_2\} = \{v, w_1\} + \{v, w_2\}.$$

We compute

$$\begin{aligned}\{v_1 + v_2, w\} &= \langle v_1 + v_2, Tw \rangle \\ &= \langle v_1, Tw \rangle + \langle v_2, Tw \rangle \\ &= \{v_1, w\} + \{v_2, w\}\end{aligned}$$

and

$$\begin{aligned}\{v, w_1 + w_2\} &= \langle v, T(w_1 + w_2) \rangle \\ &= \langle v, Tw_1 + Tw_2 \rangle \\ &= \langle v, Tw_1 \rangle + \langle v, Tw_2 \rangle \\ &= \{v, w_1\} + \{v, w_2\}\end{aligned}$$

where the second equality follows from the fact that  $T$  is a linear mapping.

Now we check that for a complex number  $c$ , we have

$$\{cv, w\} = \bar{c}\{v, w\} \text{ and } \{v, cw\} = c\{v, w\}.$$

We have

$$\begin{aligned}\{cv, w\} &= \langle cv, Tw \rangle \\ &= \bar{c}\langle v, Tw \rangle \\ &= \bar{c}\{v, w\}\end{aligned}$$

and similarly

$$\begin{aligned}\{v, cw\} &= \langle v, T(cw) \rangle \\ &= c\langle v, T(w) \rangle \\ &= c\langle v, w \rangle \\ &= c\{v, w\}.\end{aligned}$$

Finally we check whether

$$\{v, w\} = \overline{\{w, v\}}.$$

We have

$$\begin{aligned}\{v, w\} &= \langle v, Tw \rangle \\ &= \langle Tv, w \rangle \\ &= \overline{\langle w, Tv \rangle} \\ &= \overline{\{w, v\}}.\end{aligned}$$

where in the second equality we used the characterization of Hermitian operators from the **Proposition 8.6.3**. Note that this is first step of the proof where we used the hypothesis that  $T$  is Hermitian, everything before this would have "gone through" if  $T$  was just an arbitrary linear operator.

## Result

We check that this rule satisfies all the properties of a Hermitian form. Click for more details.

11. a

Suppose  $A$  is Hermitian with eigenvectors  $X_1$  and  $X_2$  with eigenvalues  $\lambda_1, \lambda_2$ , respectively. Then  $AX_1 = \lambda_1 X_1$  and  $AX_2 = \lambda_2 X_2$ . It is a part of the Spectral theorem, though it was proved in an earlier section, that  $\lambda_1$  and  $\lambda_2$  are real. Now let  $\langle, \rangle$  be the standard Hermitian product, then we have

$$\langle AX_1, X_2 \rangle = \langle \lambda_1 X_1, X_2 \rangle = \overline{\lambda_1} \langle X_1, X_2 \rangle = \lambda_1 \langle X_1, X_2 \rangle. \quad (1)$$

However, by the **Proposition 8.6.3** and noting that an operator is Hermitian if and only if its matrix with the respect to an orthonormal basis has that property -- and as  $A$  is Hermitian then it is Hermitian under the standard basis -- we have

$$\langle AX_1, X_2 \rangle = \langle X_1, AX_2 \rangle = \langle X_1, \lambda_2 X_2 \rangle = \lambda_2 \langle X_1, X_2 \rangle. \quad (2)$$

Combining (1) and (2) we have

$$\lambda_1 \langle X_1, X_2 \rangle = \lambda_2 \langle X_1, X_2 \rangle,$$

or equivalently

$$(\lambda_1 - \lambda_2) \langle X_1, X_2 \rangle = 0,$$

and as it is a hypothesis of the exercise that eigenvalues are distinct, then

$$\langle X_1, X_2 \rangle = 0,$$

which is what we needed to prove.

## Result

2 of 2

We use the proposition which gives us the equality  $\langle AX, Y \rangle = \langle X, AY \rangle$  if  $A$  is a Hermitian matrix, as well as the fact that eigenvalues of a Hermitian matrix are real, in order to show that  $\langle X_1, X_2 \rangle = 0$ , i.e.  $X_1$  and  $X_2$  are orthogonal.

## 12. a

We first find the eigenvectors of  $A$ , then normalize them and use those vectors as columns of  $P$ . Note that the Spectral theorem for Hermitian matrices -- since  $A$  is indeed Hermitian -- guarantees that eigenvectors form a basis and that this procedure diagonalizes  $A$ .

As  $A$  is a  $2 \times 2$  matrix, we can find its eigenvalues easily by finding roots of its characteristic polynomial. We compute

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 1 - \lambda & i \\ -i & 1 - \lambda \end{bmatrix} \\ &= (1 - \lambda)^2 - 1 \\ &= (1 - \lambda - 1)(1 - \lambda + 1) \\ &= -\lambda(2 - \lambda), \end{aligned}$$

which obviously has roots  $\lambda = 0$  and  $\lambda = 2$ .

Now, we have

$$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + iy \\ y - ix \end{bmatrix}.$$

Hence we want to solve

$$\begin{aligned} 2x &= x + iy \\ 2y &= y - ix, \end{aligned}$$

or equivalently

$$\begin{aligned} x - iy &= 0 \\ y + ix &= 0. \end{aligned}$$

If say  $x = 1$  then we get  $y = -i$  which together form a nontrivial solution to this system.

We also want to find eigenvector corresponding to the eigenvalue 0, so that we're looking for  $x$  and  $y$  such that

$$\begin{aligned} x + iy &= 0 \\ y - ix &= 0, \end{aligned}$$

where by picking  $x = 1$  we get  $y = i$  giving us a nontrivial solution.

We have arrived at eigenvectors

$$X_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, X_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

which are easily seen to be orthogonal and which we normalize by noting that

$$\langle X_1, X_1 \rangle = \langle X_2, X_2 \rangle = 2.$$

Now put

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} X_1 & X_2 \end{bmatrix}$$

, then

$$P^*AP = \frac{1}{2} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

## Result

We find the eigenvectors of  $A$  and normalize them in order to find that

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} X_1 & X_2 \end{bmatrix}$$

is such a unitary matrix. Click to see more details.

## 13. a

Let us describe the process we're going to go through with each of these examples. Note that each of them is symmetric, so that the Spectral theorem for symmetric operators/matrices applies. This means that the columns of  $P$  are going to be normalized (with respect to the standard dot product) eigenvectors of  $A$  (as these form an orthonormal basis), so that  $P^tAP$  is in each of these cases the diagonal matrix having the eigenvalues of  $A$  as its diagonal entries.



(a)

We have

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ 2x + y \end{bmatrix},$$

where we can immediately notice that if we choose  $x = y = 1$  we get

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix},$$

so that 3 is an eigenvalue of  $A$  associated with eigenvector

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

. As the other eigenvector must be orthogonal to

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

then it must be of the form

$$\begin{bmatrix} a \\ -a \end{bmatrix}$$

, say  $a = 1$ , then

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

so that we see that this indeed an eigenvector of  $A$  with eigenvalue  $-1$ . Now from this we obtain that

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

(b)

Similarly as previously we have

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y + z \\ x + y + z \\ x + y + z \end{bmatrix}.$$

Here it is easy to see that if  $x = y = z \neq 0$ , then this gives eigenvector

$$v_1 = [1 \ 1 \ 1]^t$$

(for the choice of  $x = y = z = 1$ ) with eigenvalue 3. Similarly we can see that if  $\lambda \neq 0$  and

$$\begin{bmatrix} x + y + z \\ x + y + z \\ x + y + z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

then  $x = y = z$ , and thus 3 is the only nonzero eigenvalue. We're looking for two choices of  $x, y, z$  such that

$$\begin{bmatrix} x + y + z \\ x + y + z \\ x + y + z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

i.e.  $x + y + z = 0$ , which is equivalent with  $z = -x - y$ . If  $v$  is in the space spanned by such vectors, then there exists  $p$  and  $q$  such that

$$v = \begin{bmatrix} p \\ q \\ -p - q \end{bmatrix} = \begin{bmatrix} p \\ 0 \\ -p \end{bmatrix} + \begin{bmatrix} 0 \\ q \\ -q \end{bmatrix} = p \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + q \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$



But note that the vectors

$$w_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^t$$

and

$$w_2 = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}^t$$

are not orthogonal. We use Gram-Schmidt in order to orthogonalize them; that is, we subtract from  $w_2$  the orthogonal projection it has onto the space spanned by  $w_1$  in order to obtain our second eigenvalue basis vector:

$$v_2 = w_2 - \frac{\langle w_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix}.$$

Now we normalize

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix}$$

in order to obtain the orthogonal eigenbasis

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix},$$

and hence our matrix  $P$  is given by

$$\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix}.$$

(c)

Let us first find the eigenvalues of  $A$  using the characteristic polynomial. We use the determinant expansion by minors in order to compute and  $3 \times 3$  determinant, whereby we obtain

$$\begin{aligned} p_A(\lambda) &= \det \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{bmatrix} \\ &= (1-\lambda) \det \begin{bmatrix} 1-\lambda & 0 \\ 0 & -\lambda \end{bmatrix} + \begin{bmatrix} 0 & 1-\lambda \\ 1 & 0 \end{bmatrix} \\ &= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda) \\ &= (1-\lambda)(-\lambda + \lambda^2 - 1) \end{aligned}$$

so that by solving  $\lambda^2 - \lambda - 1 = 0$  we obtain that the eigenvalues of  $A$  are  $1$ ,  $\frac{1}{2}(1 + \sqrt{5})$  and  $\frac{1}{2}(1 - \sqrt{5})$ . Note first that

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+z \\ y \\ x \end{bmatrix}.$$

Now if we want to find an eigenvector corresponding to  $1$ , we have to find  $x$ ,  $y$  and  $z$  not all  $0$  such that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+z \\ y \\ x \end{bmatrix}.$$

Where we note that  $x = z = 0$  by comparing the first and third entries in each vector.  $y$  can obviously be arbitrary, so we can pick the vector

$$v_1 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^t$$

In order to find an eigenvector for the eigenvalue  $\frac{1}{2}(1 + \sqrt{5})$ , we analogously have to solve

$$\left(\frac{1}{2}(1 + \sqrt{5})\right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + z \\ y \\ x \end{bmatrix},$$

from which it is immediate that  $y = 0$ . Abbreviating  $t = \frac{1}{2}(1 + \sqrt{5})$ , we're now looking for nontrivial solutions to the system

$$\begin{aligned} tx &= x + z \\ tz &= x. \end{aligned}$$

Substituting  $x = tz$  into the first equation we obtain

$$t^2 z = tz + z$$

where we can immediately notice that  $z = 1$  solves this equation as  $t^2 - t - 1 = 0$ , and hence  $x = t$ ; this gives us that an eigenvector associated to eigenvalue  $\frac{1}{2}(1 + \sqrt{5})$  is given by

$$v_2 = \begin{bmatrix} \frac{1}{2}(1 + \sqrt{5}) \\ 0 \\ 1 \end{bmatrix}.$$

Completely analogously we could find that the eigenvalue associated with  $\frac{1}{2}(1 - \sqrt{5})$  is

$$v_3 = \begin{bmatrix} \frac{1}{2}(1 - \sqrt{5}) \\ 0 \\ 1 \end{bmatrix}.$$

Now  $v_1$ ,  $v_2$  and  $v_3$  form an orthogonal basis, and it just remains to normalize them.  $v_1$  already has length 1 and we compute that

$$\langle v_2, v_2 \rangle = \frac{1}{2}(5 + \sqrt{5})$$

and

$$\langle v_3, v_3 \rangle = \frac{1}{2}(5 - \sqrt{5}).$$

and hence

$$P = \begin{bmatrix} 0 & \frac{\frac{1}{2}(1+\sqrt{5})}{\sqrt{\frac{1}{2}(5+\sqrt{5})}} & \frac{\frac{1}{2}(1-\sqrt{5})}{\sqrt{\frac{1}{2}(5-\sqrt{5})}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{\frac{1}{2}(5+\sqrt{5})}} & \frac{1}{\sqrt{\frac{1}{2}(5-\sqrt{5})}} \end{bmatrix}$$

## Result

(a)

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

, (b)

$$P = \begin{bmatrix} 1/\sqrt{3} & 1\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & -1\sqrt{2} & -1/\sqrt{6} \end{bmatrix}$$

, (c)

$$P = \begin{bmatrix} 0 & \frac{\frac{1}{2}(1+\sqrt{5})}{\sqrt{\frac{1}{2}(5+\sqrt{5})}} & \frac{\frac{1}{2}(1-\sqrt{5})}{\sqrt{\frac{1}{2}(5-\sqrt{5})}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{\frac{1}{2}(5+\sqrt{5})}} & \frac{1}{\sqrt{\frac{1}{2}(5-\sqrt{5})}} \end{bmatrix}$$

14. a

Let  $A$  be a real symmetric positive definite matrix, and suppose that one of its eigenvalues is not positive. By the Spectral theorem for symmetric operators/matrices, we have that there exists an orthogonal matrix  $P$  such that

$$P^t A P = D \quad (1)$$

and  $D = (d_{ij})$  is diagonal; it is straightforward to see that diagonal entries of  $D$  are the eigenvalues of  $A$ . Without the loss of generality (as the rows can be shuffled around via invertible matrices) say  $d_{11} \leq 0$ . Now let

$$X = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and multiply (1) from the left by  $X^t$  and from the right by  $X$ ; this turns (1) into

$$\begin{aligned} X^t P^t A P X &= X^t D X \\ (PX)^t A (PX) &= d_{11} \leq 0, \end{aligned}$$

contradicting the positive definiteness of  $A$ . Thus, all eigenvalues of  $A$  are positive.

Conversely, suppose that  $A$  is a real symmetric matrix with all its eigenvalues positive. Then, similarly as in the first part, we have an orthogonal matrix  $P$  such that

$$P^t A P = D, \quad (2)$$

where, since the eigenvalues of  $A$  are positive, all diagonal entries in the diagonal matrix  $D$  are positive. Suppose now that  $A$  is not positive definite, so there is some column vector  $X$  such that

$$X^t A X \leq 0.$$

Now as  $P$  is orthogonal, hence invertible, we can set

$$Y = P^{-1} X = P^t X.$$

But then multiplying (2) with  $Y^t$  from the left and  $Y$  from the right yields

$$(P^t X)^t P^t A P (P^t X) = Y^t D Y$$

which after some basic manipulations turns into

$$X^t A X = Y^t D Y.$$

By our assumption, the left-hand side is negative, but denoting

$$Y = [y_1 \quad \cdots \quad y_n]^t$$

we have

$$Y^t D Y = \sum_{i=1}^n d_{ii} y_i^2$$

which is a positive number because each  $d_{ii} > 0$  and  $y_i^2 \geq 0$  with at least one of them being nonzero. As we have reached a contradiction, this implies that  $A$  is indeed positive definite.

## Result

3 of 3

In both directions we used the Spectral theorem for symmetric matrices, namely the diagonalization it gives us.  
[Click to see more details.](#)

15. a

Let us first show that if  $A$  is a square matrix, then

$$\ker A = (\operatorname{im} A^*)^\perp.$$

Let  $X \in \ker A$ , then  $AX = 0$ . If  $a \in \operatorname{im} A^*$  then there exists a  $Y$  such that  $A^*Y = a$ . Now we compute

$$\begin{aligned}\langle X, a \rangle &= \langle X, A^*Y \rangle \\ &= X^*A^*Y \\ &= (AX)^*Y \\ &= 0Y = 0,\end{aligned}$$

proving that  $X$  is orthogonal to any element of  $\operatorname{im} A^*$ , i.e.

$$X \in (\operatorname{im} A^*)^\perp.$$

Now let  $Z \in (\operatorname{im} A^*)^\perp$ , then

$$\langle Z, A^*Y \rangle = 0$$

for every column vector  $Y$  of appropriate dimension such that  $A^*Y \neq 0$ . But

$$\langle Z, A^*Y \rangle = Z^*A^*Y = (AZ)^*Y,$$

and since  $(AZ)^*Y = 0$  holds for any vector  $Y$  this implies that  $(AZ)^* = 0$ , i.e.  $AZ = 0$ , which means that

$$Z \in \ker A,$$

giving us the desired equality by noting that we have obtained both inclusions.

Now let  $A$  be a normal matrix, and note that if a matrix is normal then it is immediately a square by considering the dimension of  $AA^*$  and  $A^*A$ . Observe that by using the first part, in order to show that

$$\ker A = (\operatorname{im} A)^\perp,$$

it is sufficient to show that

$$\ker A = \ker A^*,$$

for by the first part

$$\ker A^* = (\operatorname{im} A^{**})^\perp = (\operatorname{im} A)^\perp.$$

Let  $X \in \ker A$ . Then  $AX = 0$  and so  $\langle AX, AX \rangle = 0$ , but also, using the characterization which characterizes linear and normal operators (**Proposition 8.6.3**) and the normality of  $A$ , we have

$$\begin{aligned}\langle AX, AX \rangle &= \langle X, A^*AX \rangle \\ &= \langle X, AA^*X \rangle \\ &= \langle A^*X, A^*X \rangle,\end{aligned}$$

implying that  $A^*X = 0$ , i.e.  $X \in \ker A^*$ . Note that this exact line of arguments, only in reverse, shows that if  $X \in \ker A^*$  then  $X \in \ker A$ , showing that  $\ker A = \ker A^*$ .

## Result

3 of 3

For the first part we use the definitions of orthogonal space and image of a linear operator to show the equality of sets, while in the second part we show that the question reduces to proving that  $\ker A = \ker A^*$  for normal operators and show that.

First note that  $\zeta^n = e^{2\pi i} = 1$ . This is a consequence of a standard fact from complex analysis, which states that for any real number  $x$ ,

$$e^{ix} = \cos(x) + i \sin(x).$$

This identity also shows that

$$\overline{e^{ix}} = \cos(x) - i \sin(x) = \cos(-x) + i \sin(-x) = e^{-ix}.$$

Now we can compute the  $(p, q)$ th element of  $A^*A$  as

$$\sum_{j=1}^n \overline{a_{pj}} a_{qj} = \sum_{j=1}^n \frac{e^{-\frac{2\pi i}{n}(pj)}}{\sqrt{n}} \frac{e^{\frac{2\pi i}{n}(qj)}}{\sqrt{n}} = \frac{1}{n} \sum_{j=1}^n e^{\frac{2\pi i}{n}(qj-pj)} \quad (1)$$

Now in order to show that diagonal elements of  $A^*A$  are 1, note that if  $p = q$  then (1) is

$$\frac{1}{n} \sum_{j=1}^n e^{\frac{2\pi i}{n}(pj-pj)} = \frac{1}{n} \sum_{j=1}^n e^0 = \frac{1}{n} \sum_{j=1}^n 1 = 1.$$

If  $p \neq q$  then by inspecting (1) we see that if we want to show that  $A^*A = I$ , then we have to prove that

$$\sum_{j=1}^n e^{\frac{2\pi i}{n}j(q-p)} = 0. \quad (2)$$

Note now that  $q - p \neq 0$  and that  $-n < q - p < n$  (this is simply a consequence of the fact that  $q$  and  $p$  are two positive numbers between 1 and  $n$ ). Now let  $t = q - p$ , then since  $\zeta^n = 1$  we get that  $\zeta^{nt} = 1$ . We also have

$$\zeta^t = e^{\frac{2\pi i}{n}t} = \cos\left(2\pi \frac{t}{n}\right) + i \sin\left(2\pi \frac{t}{n}\right),$$

where since  $-1 < \frac{t}{n} < 1$ , i.e.  $-2\pi < 2\pi \frac{t}{n} < 2\pi$ , so that, as 0 is the only number in the interval  $(-2\pi, 2\pi)$  for which simultaneously cosine is 1 and sine is 0, then  $\zeta^t \neq 1$ . Now we can show that (2) holds, for we have

$$\begin{aligned} \sum_{j=1}^n e^{\frac{2\pi i}{n}j(q-p)} &= \sum_{j=1}^n (\zeta^t)^j \\ &= \zeta^t + (\zeta^t)^2 + \cdots + (\zeta^t)^{n-1} + 1 \\ &= \frac{1 - (\zeta^t)^n}{1 - \zeta^t} = \frac{1 - 1}{1 - \zeta^t} = 0, \end{aligned}$$

where in the second (and also the fourth) equality we used the fact that  $(\zeta^t)^n = \zeta^{tn} = 1$ , and in the third equality we used the formula for the sum of a (finite) geometric series. This completes our proof.

## Result

2 of 2

We note that  $\zeta^n = 1$  and that  $e^{ix} = \cos(x) + i \sin(x)$  for any real number, as well as show that  $\overline{e^{ix}} = e^{-ix}$ .

These facts are sufficient to demonstrate that diagonal elements of  $A^*A$  are 1, while in order to show that off-diagonal elements of  $A^*A$  are zero we prove that  $\sum_{j=1}^n \zeta^t = 0$  for any nonzero integer  $t$  such that  $-n < t < n$ .

17. a



**Given:**  $A$  and  $B$  are two Hermitian matrices such that

$$AB = BA.$$

**To Prove:** There exists a unitary matrix  $P$  such that  $P^*AP$  and  $P^*BP$  are both diagonal matrices.

**Proof:** Let us consider the principal idempotent decomposition of the matrices  $A$  and  $B$  as

$$\begin{aligned} A &= a_1A_1 + a_2A_2 + \dots + a_kA_k \\ B &= b_1B_1 + b_2B_2 + \dots + b_mB_m. \end{aligned}$$

In the above the idempotents  $A_i$  and  $B_j$  are Hermitian matrices,

for  $1 \leq i \leq k, 1 \leq j \leq m$ . Now we have  $AB = BA$ .

Then notice that

$$BA_i = A_iB, \text{ for } 1 \leq i \leq k.$$

This follows that each principal idempotents  $A_i$  of  $A$  commutes with  $B$ .

Then from the aforesaid argument it yields that each principal idempotents  $A_i$  of  $A$  commutes with each principal idempotents  $B_j$  of  $B$ .

This is

$$A_iB_j = B_jA_i, \text{ for all } 1 \leq i \leq k, 1 \leq j \leq m.$$

Let us now consider the matrix

$$C = \sum_{i,j} c_{i,j} A_i B_j, \text{ where } c_{i,j} \text{ are all distinct.}$$

Since linearity holds in the vector space of Hermitian matrices we have  $C$  is an Hermitian Matrix.

Let us now define the real polynomials  $f(x)$  and  $g(x)$  by the assignment

$$f(c_{i,j}) = a_i \text{ and } g(c_{i,j}) = b_j \text{ for all } i, j.$$

Now we have

$$\begin{aligned} C^2 &= \left( \sum_{i,j} c_{i,j} A_i B_j \right) \left( \sum_{i,j} c_{i,j} A_i B_j \right) \\ &= \left( \sum_{i,j} c_{i,j}^2 A_i B_j \right), \text{ since } A_i B_j = B_j A_i. \end{aligned}$$

By the similar argument it yield's that

$$C^m = \left( \sum_{i,j} c_{i,j}^n A_i B_j \right). \quad (1)$$

Now notice that

$$\begin{aligned} f(C) &= f\left(\sum_{i,j} c_{i,j} A_i B_j\right) \\ &= \sum_{i,j} f(c_{i,j}) A_i B_j, \text{ by (1)} \\ &= \sum_{i,j} a_i A_i B_j, \text{ by the definition of } f \\ &= \sum_j \left(\sum_i a_i A_i\right) B_j \\ &= \sum_j A B_j \\ &= A \sum_j B_j \\ &= A, \text{ since } \sum_j B_j = 1. \end{aligned}$$

Similarly we have

$$\begin{aligned} g(C) &= g\left(\sum_{i,j} c_{i,j} A_i B_j\right) \\ &= \sum_{i,j} g(c_{i,j}) A_i B_j \\ &= \sum_{i,j} b_j A_i B_j, \text{ by the definition of } g \\ &= \sum_i \left(\sum_j b_j B_j\right) A_i \\ &= \sum_i B A_i \\ &= B \sum_i A_i \\ &= B, \text{ since } \sum_i A_i = 1. \end{aligned}$$

This shows that  $A$  and  $B$  are polynomials with real coefficients in a common Hermitian Matrix  $C$ .

Now since  $C$  is a Hermitian Matrix there exists a Unitary Matrix  $P$  such that  $P * CP$  is a diagonal matrix, say  $D$ .

Now notice that

$$\begin{aligned} P * AP &= f(P * CP) = f(D) \\ P * BP &= g(P * CP) = g(D). \end{aligned}$$

This follows that both the matrices  $P * AP$  and  $P * BP$  are diagonal.

This completes the proof.

## Result

4 of 4

First we show that if  $A$  and  $B$  commutes then both are polynomials with real coefficients in a common Hermitian Matrix  $C$  and then proves the result by using it.

18. a

Note that by the **Exercise 6.14**, all of eigenvalues of  $A$  are positive. Therefore, applying the Spectral theorem for symmetric operators/matrices to  $A$  we find there exists an orthogonal matrix  $Q$  such that

$$Q^t A Q = D \quad (1)$$

where  $D$  is a diagonal matrix with diagonal entries being the eigenvalues of  $A$ , and we write  $D = (d_{ij})$ . As  $Q$  is orthogonal and hence  $Q Q^t = Q^t Q = I$ , we can multiply (1) from the left by  $Q$  and from the right by  $Q^t$  to obtain

$$A = Q D Q^t \quad (2)$$

Note that since  $D$  is a diagonal matrix and  $d_{ii} > 0$  for all  $i = 1, \dots, n$ , then

$$D = \sqrt{D} \sqrt{D} = (\sqrt{d_{ij}})(\sqrt{d_{ij}}),$$

where  $\sqrt{D}$  is a diagonal  $n \times n$  matrix such that  $(i, i)$ th entry is  $\sqrt{d_{ii}}$ . Now we substitute this into (2) to get

$$A = (Q \sqrt{D})(\sqrt{D} Q^t) = (Q \sqrt{D})(\sqrt{D}^t Q^t) = (Q \sqrt{D})(Q \sqrt{D})^t \quad (3)$$

where we used the fact that  $\sqrt{D} = \sqrt{D}^t$  because  $\sqrt{D}$  is a diagonal matrix.

Now putting  $P = (Q \sqrt{D})^t$  the desired result follows.

**Result**

2 of 2

We use the previous exercise which states that all of eigenvalues of a positive definite symmetric real matrix are positive which via the Spectral theorem then yields a way to explicitly write down a form for  $P$  such that  $A = P^t P$ .  
[Click to see more details.](#)

19. a

Let  $S$  be the  $n \times n$  cyclic shift operator. Since  $S$  is a real matrix, we have  $S^* = S^t$  and therefore in order to show that  $S$  is unitary we have to show that

$$S S^t = I.$$

We show this by a direct calculation; specifically, we determine  $(i, j)$ th entry in  $S S^t$ . Suppose first  $i < n$  and  $j < n$ . Then  $(i, j)$ th entry is a dot product of  $i$ th row of  $S$  -- and it is such that every but  $(i + 1)$ th entry is zero, while  $(i + 1)$ th entry is 1 -- with  $j$ th column of  $S^t$ , which is the same as  $j$ th row of  $S^t$ , and now it immediately follows from the description of  $i$ th row of  $S$  that  $(i, j)$ th entry is 0 except if  $i = j$ , in which case it is  $1 \cdot 1 = 1$ .

If  $i < n$  and  $j = n$  then  $(i, n)$ th entry is a dot product of a row with only  $(i + 1)$ th entry equal to 0 and the  $n$ th column of  $S^t$  -- that being the  $n$ th row of  $S$ , which has 1 only in the first position, and thus the dot product is always 0 as  $i + 1 > 1$ . If  $i = n$  and  $j < n$  then an analogous argument shows that  $(n, j)$ th entry is 0.

Finally  $(n, n)$ th entry is a dot product of  $(1, 0, \dots, 0)$  (the  $n$  row in  $S$ ) and  $(1, 0, \dots, 0)$  ( $n$ th column in  $S^t$ , which is the same as  $n$ th row in  $S$ ), and therefore it's equal to 1.

In order to find the diagonalization of  $S$ , we find its eigenvalues. First note that the name "cyclic shift operator" from the property of  $S$  that

$$S \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 \\ \vdots \\ x_n \\ x_1 \end{bmatrix}.$$

(This is straightforward to obtain by direct computation.) Suppose now  $\lambda$  is an eigenvalue of  $S$ , then we have

$$\begin{bmatrix} x_2 \\ \vdots \\ x_n \\ x_1 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

which gives us  $n$  equalities

$$\begin{aligned} x_2 &= \lambda x_1 \\ x_3 &= \lambda x_2 \\ &\vdots \\ x_n &= \lambda x_{n-1} \\ x_1 &= \lambda x_n. \end{aligned}$$

As eigenvectors are by definition nonzero, and since if any of those numbers is 0 then each of them is 0, we can assume they're all nonzero. Then, by substituting first the first equation into the second, then the second into the third, etc., we obtain

$$x_1 = \lambda^n x_1,$$

and since  $x_1 \neq 0$ , this implies that

$$\lambda^n = 1.$$

Furthermore, we can see that if  $\lambda^n = 1$  then it is an eigenvalue, as we can go in the "opposite direction". It is a basic fact (from complex analysis, but also used/proved elsewhere; see the Exercise 16 for an idea how to prove it assuming Euler's identity) that  $\lambda^n = 1$  has  $n$  distinct solutions (called  $n$ th roots of unity) given by  $1, \zeta, \zeta^2, \dots, \zeta^{n-1}$ , where  $\zeta = e^{\frac{2\pi i}{n}}$ . Hence the diagonalization of  $S$  is a diagonal matrix with entries  $1, \zeta, \zeta^2, \dots, \zeta^{n-1}$ .

## Result

3 of 3

First we show that  $S$  is unitary by direct computation, and then we find its diagonalization by determining that its eigenvalues are all the  $n$ th roots of unity. Click to see more details.

20. a

Recall that a matrix  $A$  is normal if and only if it commutes with its adjoint, i.e.  $AA^* = A^*A$ . Now let  $C$  be the circulant matrix. We're going to show that  $(i, j)$ th entry, for arbitrary  $i, j$ , of  $CC^*$  is the same as that of  $C^*C$ .

## Step 2

2 of 4

First let us examine the  $(i, j)$ th entry of  $CC^*$ . It is the dot product of the  $i$ th row of  $C$  and conjugated  $j$ th row of  $C$  (since that is what the  $j$ th column of  $C^*$  is). First, if  $i = j$  then we see that  $i$ th row contains all the numbers  $c_0, \dots, c_n$ , and hence the diagonal entry will be equal to

$$\sum_{i=0}^n c_i \overline{c_i}.$$

Now let  $i \neq j$ , inspecting further we see that its  $i$ th row of  $C$  is (written as a row matrix)

$$[c_{n-i+2} \quad \cdots \quad c_n \quad c_0 \quad \cdots \quad c_{n-i}]$$

(With convention that  $c_i$  is an empty string for  $i > n$  or  $i < 0$  so we don't have to treat first and last row separately.) Note that since  $CC^*$  is a Hermitian matrix then we can assume that  $i < j$ , and hence  $j = i + k$  for some positive  $k$ .

Now we have that the  $(i, i + k)$ th entry of  $CC^*$  is equal to

$$c_{n-i+2} \overline{c_{n-i-k+2}} + \cdots + c_0 \overline{c_{n-k+2}} + \cdots + c_k \overline{c_0} + \cdots + c_{n-i} \overline{c_{n-i-k}}. \quad (1)$$

Now we examine the  $(i, j)$ th entry of  $C^*C$ . It is the dot product of  $i$ th row of  $C^*$  -- which is the conjugated  $i$ th column of  $C$  -- and  $j$ th column of  $C$ . We see that every column contains each of the number  $c_0, \dots, c_n$  so that the dot product of a conjugated column with itself without conjugation -- which is what happens in the case when  $i = j$  -- is

$$\sum_{i=0}^n \overline{c_i} c_i = \sum_{i=0}^n c_i \overline{c_i}.$$

For the case  $i \neq j$  note that the  $i$ th column of  $C$  is (written as a row matrix)

$$[c_{i-1} \quad c_{i-2} \quad \cdots \quad c_0 \quad c_n \quad c_{n-1} \quad \cdots \quad c_i]$$

(with the convention that  $c_{n+1} = c_0$  so we don't have to treat the last column separately.)

Now let again  $i < j$ ,  $j = i + k$ , with the same reasoning as in the previous case. Then  $(i, i + k)$ th entry is equal to

$$\overline{c_{i-1}} c_{i+k-1} + \cdots + \overline{c_{n-i-k}} c_{n-i} + \cdots + \overline{c_0} c_k + \cdots + \overline{c_{n-k+2}} c_0 + \cdots + \overline{c_{n-i-k+2}} c_{n-i+2}. \quad (2)$$

By carefully comparing (1) and (2), as well as the way that those sums arise, we note their equality.

## Result

4 of 4

We directly compute the  $(i, j)$ th entry if  $CC^*$  and  $C^*C$  and note their equality. Click to see more details.

21. a



Recall that by the Spectral Theorem for Normal operators/matrices we have unitary matrix  $P$  such that

$$P^*AP = D, \quad (1)$$

where  $D$  is diagonal with its entries being eigenvalues of  $A$ . By conjugating both sides of (1) we get

$$P^*A^*P = D^*.$$

Now if we know that  $D^* = D$ , then

$$P^*AP = P^*A^*P$$

and by multiplying by  $P$  from the left and  $P^*$  from the left and recalling that since  $P$  is unitary then  $P^*P = P^*P = I$ , we have

$$A^* = A.$$

When do we have  $D^* = D$ ? Recalling that the adjoint is conjugate transpose, then if  $D = (d_{ij})$  we have  $D^* = (\overline{d_{ji}})$ , and thus this equality holds if and only if  $d_{ii} = \overline{d_{ii}}$  for any  $i$ , which is equivalent with  $d_{ii}$  being a real number. Therefore, a normal matrix with real eigenvalues is Hermitian.

Similarly as in the first part, we have

$$P^*AP = D \text{ and } P^*A^*P = D^*,$$

where by multiplying them we get

$$(P^*AP)(P^*A^*P) = P^*AA^*P = DD^*.$$

Note now that if  $D = (d_{ij})$  then  $DD^* = (d_{ij}\overline{d_{ij}})$  and hence if we want to conclude  $AA^* = I$  (which means that  $A$  is unitary) we must have  $DD^* = I$ , i.e.  $\lambda\overline{\lambda} = 1$  for any eigenvalue  $\lambda$  of  $A$ . Observe that  $\lambda\overline{\lambda} = |\lambda|$ .

## Result

3 of 3

We find that a normal matrix is Hermitian if its eigenvalues are real, and that it is unitary if its eigenvalues have norm 1, i.e. be on the unit circle in the complex plane. Click to see more details.

## 22. a

First we prove analogues of **Proposition 8.6.3**, **Proposition 8.6.4** and **Theorem 8.6.5** for symmetric operators on a Euclidean space. In order to introduce the analogues, for a linear operator  $T$ , with matrix  $A$  is some orthonormal basis, we define its transpose operator  $T^t$  to be the operator which has the matrix  $A^t$  with respect to that same basis. Note that the proof of its well-definedness would be completely analogous to that of adjoint operators, only that we would replace the word "unitary" with "orthogonal" and matrix adjoints with matrix transpositions.

We define an operator  $T$  to be symmetric if  $T = T^t$ , or equivalently if its corresponding matrix is symmetric.



(Analogue of **Proposition 8.6.3**)

We prove that if  $T$  is a symmetric operator and  $\langle \cdot, \cdot \rangle$  is a form on a Euclidean space  $V$ , then for all  $v, w$  in  $V$  we have

$$\langle Tv, w \rangle = \langle v, T^t w \rangle \quad (1)$$

and

$$\langle Tv, Tw \rangle = \langle T^t v, T^t w \rangle. \quad (2)$$

First fix an orthonormal basis, let  $A$  be the matrix of  $T$  with the respect to it, and let  $X$  and  $Y$  be the column vectors of  $v$  and  $w$ , respectively. Then we have

$$\langle Tv, w \rangle = (AX)^t Y = X^t (A^t Y) = \langle v, T^t w \rangle$$

showing (1). Completely analogously we could show that  $\langle v, Tw \rangle = \langle T^t v, w \rangle$ , and

$$\langle Tv, Tw \rangle = (AX)^t (AY) = (A^t X)^t (A^t Y) = \langle T^t v, T^t w \rangle,$$

where in the third equation we used the symmetry of  $A$ , i.e.  $A = A^t$ , showing (2).

We also need to prove -- we need this in the proof of Spectral theorem since it implies that a restriction of a symmetric operator to an invariant subspace is again symmetric -- that if  $\langle Tv, w \rangle = \langle v, Tw \rangle$ , then  $T$  is symmetric.

By (1) we have  $\langle Tv, w \rangle = \langle v, T^t w \rangle$ , i.e.

$$\langle v, T^t w \rangle = \langle v, Tw \rangle$$

for all  $v$  and  $w$ , and therefore  $T^t w = Tw$  for all  $w$ , implying that  $T = T^t$  which means that  $T$  is symmetric.

(Analogue of **Proposition 8.6.4**) Let  $T$  be a symmetric operator on a Euclidean space  $V$ ,  $W$  a subspace of  $V$ . If  $W$  is  $T$ -invariant, then the orthogonal space  $W^\perp$  is  $T^t$ -invariant, and if  $W$  is  $T^t$  invariant then  $W^\perp$  is  $T$  invariant. Note that now that we have proved an analogue of **Proposition 8.6.3**, a basically word-for-word (replacing adjoints with transposes) proof from the text works for our case.

#### Step 4

4 of 7

(Analogue of **Theorem 8.6.5**) If  $T$  is a symmetric operator on a Euclidean space and  $v$  is an eigenvector of  $T$  with eigenvalue  $\lambda$ , then  $v$  is also an eigenvector of  $T^t$  with eigenvalue  $\lambda$ . This follows easily from the observation that  $T^t = T$ .

Now we are ready to proceed to the proof of the Spectral theorem. As it has been noted in the short proof sketch in the text, it was proved in an earlier section that the eigenvalues of a real symmetric matrix are real numbers. Then it quickly follows that since an operator and its associated matrix share eigenvalues, and since an operator is symmetric if and only if its matrix is symmetric, hence the eigenvalues of a symmetric operator are real.

We use this to prove that if  $T$  is a symmetric operator on a Euclidean space  $V$ , then there is an orthonormal basis of  $V$  consisting of eigenvectors of  $T$ .

First we choose an eigenvector  $v_1$  for  $T$ , which we can normalize so that  $|v_1| = 1$ . By our analogue of **Theorem 8.6.5** we have that  $v_1$  is also an eigenvalue for  $T^t$  (actually, this is obvious just as the **Theorem 8.6.5** is obvious in this case).

Hence the space  $W$  spanned by  $v_1$  is  $T^t$ -invariant, and by our analogue of **Proposition 8.6.4**, this means that  $W^\perp$  is  $T$ -invariant. Since we are in a Euclidean space and hence the form is nondegenerate on it, then

$$V = W \oplus W^\perp.$$

As have noted before, the restriction of  $T$  to  $W^\perp$  is a symmetric operator. If  $W^\perp$  is of dimension 1 then we are done as the invariance of  $W^\perp$  under  $T$  would imply existence of an eigenvector.

In the general case with dimension of  $W^\perp$  greater than 1, we note that since the eigenvalues of  $T$  are real, then if we have  $AX = \lambda X$  for real  $\lambda$ , if we decompose  $X$  into real and imaginary parts, i.e.  $X = X_1 + iX_2$ , then  $X_1$  and  $X_2$  are real matrices and  $AX_1 = \lambda X_1$  and  $AX_2 = \lambda X_2$ , i.e. an operator/matrix with real eigenvalues has a real eigenvector.

Hence  $W^\perp$ , being a subspace of a Euclidean space, contains at least one eigenvector, say  $v_2$ , which allows us to iteratively extended the basis by considering the space  $W_2$  spanned by  $v_2$ , in which case  $W^\perp = W_2 \oplus W_2^\perp$ , etc., in order to obtain an orthonormal basis of eigenvalues,  $v_1, \dots, v_n$ .

The matrix version now again follows analogously as for the normal operators, except that the change of basis is orthogonal rather than unitary.

## Result

7 of 7

We define the transpose operator  $T^t$  for a linear operator  $T$  on a Euclidean space, prove various proposition analogous to those in the text, and then use them to give a proof of the Spectral theorem for symmetric operators.

## Section 7

1. a

Using the notation of the section on conics and quadrics, we have that the matrix  $A$  associated with this quadratic equation is

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 3 & 0 & 1 \end{bmatrix}.$$

A short computation on a computer-algebra system shows that eigenvalues and hence eigenvectors of  $A$  are beyond cumbersome, so we opt for a non-orthogonal change of basis. We apply an algorithm (see Exercise 14 in the Section 4 for details) for diagonalizing via row and column operations in order to find a change of basis which diagonalizes  $A$ . We have

$$\begin{aligned} \left( \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) &\sim \left( \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -4 & -2 & -2 & 1 & 0 \\ 0 & -2 & 0 & -1 & 0 & 1 \end{array} \right) \\ &\sim \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -4 & -2 & -2 & 1 & 0 \\ 0 & -2 & 0 & -1 & 0 & 1 \end{array} \right) \\ &\sim \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -4 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 1 \end{array} \right) \\ &\sim \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -4 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 1 \end{array} \right), \end{aligned}$$

so that

$$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}^t \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Changing  $B$  to that basis we also get

$$B' = \begin{bmatrix} 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 1 \end{bmatrix}.$$

Hence our quadric is transformed into

$$f(x, y, z) = x^2 - 4y^2 + z^2 + 3x + 6y + z - 6,$$

where we "complete the squares" with substitutions  $x' = x - 3/2$ ,  $y' = y + 6/8$ ,  $z' = z - 1/2$ , so that we obtain (dropping primes)

$$f(x, y, z) = x^2 - 4y^2 + z^2 - \frac{25}{4}.$$

By the theorem given in the text which classifies the quadrics (**Theorem 8.7.14**) we see that this is a one-sheeted hyperboloid.

## Result

3 of 3

We find it is a one-sheeted hyperboloid. Click to see more details.

2. a

Consider the function  $f$  given by

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c$$

The loci  $f = 0$  represent the various conic-sections in the plane  $\mathbb{R}^2$

Consider the function given by

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c$$

Let  $f(x_1, x_2) = 0$  represent the equation of an ellipse.

Let  $(h, k)$  denote the center of the ellipse

Now at the center of the ellipse the following condition holds true

$$\left(\frac{df}{dx_1}\right)_{x_1=h} = 0 \text{ and } \left(\frac{df}{dx_2}\right)_{x_2=k} = 0$$

So evaluate  $\frac{df}{dx_1}$  and  $\frac{df}{dx_2}$

$$\begin{aligned} \frac{df}{dx_1} &= \frac{d(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c)}{dx_1} \\ &= 2a_{11}x_1 + 2a_{12}x_2 + b_1 \end{aligned}$$

$$\begin{aligned} \frac{df}{dx_2} &= \frac{d(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c)}{dx_2} \\ &= 2a_{22}x_2 + 2a_{12}x_1 + b_2 \end{aligned}$$

Since  $\left(\frac{df}{dx_1}\right)_{x_1=h} = 0$  and  $\left(\frac{df}{dx_2}\right)_{x_2=k} = 0$ . This gives the system of 2 linear equation in 2 variables which are as follows

$$2a_{11}h + 2a_{12}k + b_1 = 0$$

$$2a_{22}k + 2a_{12}h + b_2 = 0$$

Use the method of elimination to solve the above system of equation

Multiply the first equation by  $(a_{12}/a_{11})$  and second by 1

$$\begin{aligned} 2a_{12}h + 2\left(\frac{a_{12}}{a_{11}}\right)^2k &= \frac{-a_{12}b_1}{a_{11}} \\ 2a_{12}h + 2a_{22}k &= -b_2 \end{aligned}$$

Now subtract both the equations to eliminate the variable  $x_1$

$$\begin{aligned} 2k\left(\frac{(a_{12})^2}{a_{11}} - a_{22}\right) &= b_2 - \frac{b_1a_{12}}{a_{11}} \\ 2k\left(\frac{(a_{12})^2}{a_{11}} - a_{11}a_{22}\right) &= \frac{b_2a_{11} - b_1a_{12}}{a_{11}} \\ k &= \frac{1}{2}\left(\frac{b_2a_{11} - b_1a_{12}}{(a_{12})^2 - a_{11}a_{22}}\right) \end{aligned}$$

Now multiply the second equation by  $(a_{12}/a_{22})$  and first by 1

$$\begin{aligned} 2a_{11}h + 2a_{12}k + b_1 &= 0 \\ 2\left(\frac{(a_{12})^2}{a_{22}}\right)h + 2a_{12}k &= \frac{-a_{12}b_2}{a_{22}} \end{aligned}$$

Now subtract both the equation to eliminate the variable  $x_2$

$$\begin{aligned} 2h\left(\frac{(a_{12})^2}{a_{22}} - a_{11}\right) &= b_1 - \frac{b_2a_{12}}{a_{22}} \\ 2h\left(\frac{(a_{12})^2}{a_{22}} - a_{11}a_{22}\right) &= \frac{b_1a_{22} - b_2a_{12}}{a_{22}} \\ h &= \frac{1}{2}\left(\frac{b_1a_{22} - b_2a_{12}}{(a_{12})^2 - a_{11}a_{22}}\right) \end{aligned}$$

Define a new map  $g(x_1, x_2)$  which is obtained by translating the function  $f(x_1, x_2)$  to the center  $(h, k)$  given by

$$(h, k) = \left( \frac{1}{2} \left( \frac{b_1 a_{22} - b_2 a_{12}}{(a_{12})^2 - a_{11} a_{22}} \right), \frac{1}{2} \left( \frac{b_2 a_{11} - b_1 a_{12}}{(a_{12})^2 - a_{11} a_{22}} \right) \right)$$

So,  $g(x_1, x_2) = f(x_1 + h, x_2 + k)$

Under this translation the function  $g(x_1, x_2)$  attains the form

$$g(x_1, x_2) = Ax_1^2 + 2Bx_1x_2 + Cx_2^2 + F = 0, \text{ where the coefficients can be evaluated}$$

**Therefore, one can translate the quadratic equation as mentioned in the above solution.**

### 3. a

Suppose that

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + b_1x + b_2y + c$$

and so  $f = 0$  is a conic. First recall that a circle is an ellipse (as given by the classification of ellipses in the **Theorem 8.7.5**) with coefficients  $a_{11}$  and  $a_{22}$  being equal. Hence if the conic is a circle, we must have that

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$$

has a diagonalization of the form

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix},$$

$a > 0$  (greater than zero WLOG, otherwise we just multiply the whole conic with  $-1$ ). But this means there exist an orthogonal  $P$  such that

$$P^t AP = aI,$$

whereupon multiplying by  $P$  from the left and  $P^t$  from the right, and noting that scalar multiplication and multiplication of  $I$  with any matrix commute, we obtain

$$A = aI,$$

and therefore

$$a_{11} = a_{22} = a > 0 \text{ and } a_{12} = 0; \quad (1)$$

now we have

$$f(x, y) = ax^2 + ay^2 + b_1x + b_2y + c.$$

We introduce the substitutions

$$x' = x - \frac{b_1}{2a} \text{ and } y' = y - \frac{b_2}{2a} \quad (2)$$

(we shift the potential circle to the center with this transformation) in order to obtain (after dropping primes)

$$f(x, y) = ax^2 + ay^2 + \frac{b_1^2}{4a^2} + \frac{b_2^2}{4a^2} + c,$$

and so for  $f = 0$  to be circle we must also have

$$\frac{b_1^2}{4a^2} + \frac{b_2^2}{4a^2} + c < 0,$$

or equivalently

$$b_1^2 + b_2^2 + 4ca^2 < 0. \quad (3)$$

(This actually ensures that the conic is nondegenerate.)



Hence these are necessary conditions, and let us comment briefly how to see that they are also sufficient. Suppose there was a conic satisfying (1) and (3). Then after performing the change of coordinates (2) (which only moves the conic), and abbreviating the left-hand side of (3) as  $k$ , we know that its equation is

$$x^2 + y^2 = \frac{-t}{a}.$$

But since  $\frac{-t}{a} > 0$  we know that this describes a circle on the plane.

## Result

3 of 3

A necessary and sufficient condition for a conic  $f = 0$  with  $f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + b_1x + b_2y + c$  to be a circle is given by  $a_{11} = a_{22} \neq 0$ , and  $b_1^2 + b_2^2 + 4ca_{11}^2 < 0$  if  $a_{11} > 0$  and  $b_1^2 + b_2^2 + 4ca_{11}^2 > 0$  if  $a_{11} < 0$ . Click for more details.

## 4. a

Let  $f$  be a quadric, i.e.

$$f(x_1, x_2, x_3) = X^t A X + B X + c,$$

as described in (8.7.12). The case when  $B = 0$  and  $c = 0$  is already determined in the text to be a double cone, that is a union of lines through origin.

Inspecting **Theorem 8.7.14**, we see that the following classes are not categorized there:

- (a)  $a_{11}x_1^2 + a_{22}x_2^2 - 1 = 0$ ,
- (b)  $a_{11}x_1^2 - a_{22}x_2^2 - 1 = 0$ ,
- (c)  $a_{11}x_1^2 - x_2 = 0$ ,
- (d)  $a_{11}x_1^2 + a_{22}x_2^2 - x_3^2 = 0$ ,
- (e)  $a_{11}x_1^2 + a_{22}x_2^2 + x_3^2 = 0$ ,

where we have ignored the cases when coefficient next to quadratic terms are 0, as these are just planes/lines, as well as some cases which reduce to the listed ones through substitutions, those that form an empty set or set with only one point, etc.

We see that (a),  $a_{11}x_1^2 + a_{22}x_2^2 - 1 = 0$ , is a union of ellipses, i.e. an ellipse on each plane  $x_3 = r \in \mathbb{R}$ . This shape is called **elliptic cylinder**.

Similarly, (b),  $a_{11}x_1^2 - a_{22}x_2^2 - 1 = 0$ , is such a union of hyperbolas. It is called **hyperbolic cylinder**.

In (c) too we can recognize an equation for a parabola. As it does not depend on  $x_3$ , it is again a union of parabolas, and hence called **parabolic cylinder**.

The (d) and (e) were already classified in the text.

## Result

3 of 3

In addition to the degenerate quadrics already classified in the text, we also find there are so-called elliptic cylinders, hyperbolic cylinders, and parabolic cylinders. Click for more details.



## Section 8

1. a

Since  $A$  is skew-symmetric then  $A^t = -A$ , and hence  $A = -A^t$ . Now we compute

$$A^2 = AA = (-A^t)(-A^t) = A^t A^t = (A^t)^2 = (A^2)^t,$$

showing that  $A^2$  is symmetric. Note that we could interchange transposition and squaring because  $(A^2)^t = (AA)^t = A^t A^t = (A^t)^2$ .

Now we want to prove that  $A^2$  is negative definite. Let  $X$  be a nonzero column vector, then since we have  $A = -A^t$ , we have

$$\begin{aligned} X^t A^2 X &= X^t A A X \\ &= X^t (-A^t) A X \\ &= -X^t A^t A X \\ &= -(AX)^t A X \end{aligned}$$

where we note that since  $AX$  is a column vector, say

$$AX = [a_1 \quad \cdots \quad a_n]$$

, then  $(AX)^t A X = \sum_{i=1}^n a_i^2$ , and hence since  $a_i$  are real numbers we have

$$(AX)^t A X \geq 0$$

and hence

$$-(AX)^t A X = X^t A^2 X \leq 0.$$

To see that the inequality is strict, note that for equality to hold we would have to have  $AX = 0$  for a nonzero  $X$ ; but as  $A$  is invertible this is impossible. This completes our proof.

### Result

3 of

Symmetry of  $A^2$  is straightforward from noting that since  $A$  is skew-symmetric then  $A = -A^t$ , and substituting this for  $A$ 's. Negative definiteness also follows from this observation and substitution. Click to see more details.

2. a

Since a real skew-symmetric form on  $W$  is nondegenerate, it follows that the dimension of  $W$  is an even integer (this is **Corollary 8.8.8**). Hence, again by the results of that section, we can choose a basis  $w_1, \dots, w_{2n}$  such that the matrix of the form is made up of diagonal blocks of the form

$$\Sigma = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Therefore, it follows directly from the form of the matrix of the form that

$$\langle w_i, w_{i-1} \rangle = 1 \text{ if } i \text{ is even,} \quad (1)$$

$$\langle w_i, w_{i+1} \rangle = -1 \text{ if } i \text{ is odd,} \quad (2)$$

and  $\langle w_i, w_j \rangle = 0$  in any other case.

Now let  $\pi : V \rightarrow W$  be an orthogonal projection, where we write

$$\pi(v) = \sum_{i=1}^{2n} t_i w_i.$$

We want to determine the form of  $t_i$ . In order to do this, note first that as  $\pi$  is an orthogonal projection, then we have that for any vector  $v \in V$ ,  $v - \pi(v)$  is orthogonal to  $W$ , i.e.

$$\langle w_i, v - \pi(v) \rangle = 0$$

for  $i = 1, \dots, 2n$ . Therefore, by the bilinearity of the form we have

$$\langle w_i, v - \pi(v) \rangle = \langle w_i, v \rangle - \sum_{j=1}^{2n} \langle w_i, w_j \rangle t_j. \quad (3)$$

Note that in (3) we have  $i$  fixed and  $j$  spanning from 1 to  $2n$ ; therefore, we need to consider two cases, one when  $i$  is even and another when  $i$  is odd. If  $i$  is even, then evaluating the sum in (3) and using the fact that  $v - \pi(v)$  is orthogonal to  $W$  we get

$$0 = \langle w_i, v \rangle - \langle w_i, w_{i-1} \rangle t_{i-1},$$

where recalling (1) we have

$$0 = \langle w_i, v \rangle - t_{i-1},$$

implying that

$$t_{i-1} = \langle w_i, v \rangle.$$

Completely analogously it would follow that if  $i$  is odd, then  $t_{i+1} = -\langle w_i, v \rangle$ .

Putting these together we obtain

$$\pi(v) = \sum_{i=1}^{2n} (-1)^{i+1} \langle w_i, v \rangle w_i.$$

(Exponent  $i+1$  comes from the fact that we want the sign to be positive when  $i-1$  is even, i.e. when  $i$  is odd, and negative when  $i+1$  is odd, i.e.  $i$  is even.)

## Result

3 of 3

Let  $\pi : V \rightarrow W$  be the orthogonal projection,  $v \in V$ , then we show that the orthogonal projection is of the form

$$\pi(v) = \sum_{i=1}^{2n} (-1)^{i+1} \langle w_i, v \rangle w_i.$$

[Click to see more details.](#)

## 3. a

We first show that  $I + S$  is invertible. Note that if  $AB$  for some square matrices is invertible, then so is  $A$ ; this is a consequence of the multiplicativity of the determinant, for  $\det(AB) = \det A \det B$ , and so  $\det(AB) \neq 0$  implies  $\det A \neq 0$ . Therefore it suffices to show that

$$(I + S)(I - S) = I^2 - S + S - S^2 = I - S^2 \quad (1)$$

is invertible. First note that by the first exercise in this section,  $S^2$  is symmetric (we did **not** use the hypothesis of invertibility in that part of the proof). Furthermore, the second part of my proof of that exercise also shows that  $S^2$  is negative semidefinite. Now reasoning analogous to that of **Exercise 14** in the section on Spectral theorem shows that all eigenvalues of a negative semidefinite matrix are  $\leq 0$ .

Applying the Spectral theorem for symmetric operators to  $S^2$  we obtain that there exists an orthogonal matrix  $P$  such that

$$P^t S^2 P = D,$$

$D$  is diagonal, and by the characterization of eigenvalues of  $S^2$  we have that all its entries are nonpositive. We also have

$$P^t(I - S^2)P = P^t IP - P^t S^2 P = I - D,$$

where recalling that all entries of  $D$  are nonpositive, we obtain that  $I - D$  is a diagonal matrix with all positive and therefore nonzero diagonal entries. This shows that  $I - D$  is invertible, say it has an inverse  $C$ , then

$$(P^t(I - S^2)P)C = I,$$

where by first multiplying by  $P$  from the left and then by  $P^t$  from the right, and noting the orthogonality of  $P$ , we obtain

$$(I - S^2)PCP^* = I.$$

Now, recalling (I) we see that we're done.

Let us now show that  $(I - S)(I + S)^{-1}$  is orthogonal. We are going to be using this shortly, so let us first prove that  $I + S$  and  $I - S$  commute. This is straightforward by direct computation, for

$$(I + S)(I - S) = I^2 - S + S - S^2 = I - S^2$$

and

$$(I - S)(I + S) = I^2 + S - S - S^2 = I - S^2.$$

Now we compute

$$\begin{aligned} [(I - S)(I + S)^{-1}]^t &= [(I + S)^{-1}]^t (I - S)^t \\ &= [(I + S)^t]^{-1} (I - S)^t \\ &= (I - S)^{-1} (I + S) \end{aligned}$$

where note that by skew-symmetry of  $S$  and the property of transposition that  $(A + B)^t = A^t + B^t$ , we have  $(I \pm S)^t = I \mp S$ . Now we multiply

$$\begin{aligned} [(I - S)(I + S)^{-1}]^t (I - S)(I + S)^{-1} &= (I - S)^{-1} (I + S) (I - S)(I + S)^{-1} \\ &= (I - S)^{-1} (I - S) (I + S)^{-1} \\ &= I, \end{aligned}$$

where the first equality follows from what we just proved, and the second equality follows from the fact that  $I + S$  and  $I - S$  commute. This completes the proof of orthogonality of  $(I - S)(I + S)^{-1}$ .

## Result

4 of 4

For the first part we use the first exercise as well as the characterization of eigenvalues of a negative semidefinite symmetric matrix and the spectral theorem. Second part of the exercise follows from direct computation and observation that  $I + S$  and  $I - S$  commute. Click for more details.

4. a

**Given:**  $A$  is a real skew-symmetric matrix.

**To Prove:**

(a)  $\det(A) \geq 0$ .

(b) If  $A$  has integer entries, then  $\det(A)$  is a perfect square.

**Proof:** First we start with a lemma.

Lemma: A skew-symmetric matrix of odd order have determinant zero.

Proof of the lemma: Let  $A$  be a skew-symmetric matrix of odd order.

Now since  $A$  is a skew-symmetric matrix we have

$$A^T = -A, \text{ where } A^T \text{ denote the transpose of } A.$$

Also note that for any matrix  $A$

$$\det(A) = \det(A^T).$$

Then from the above argument it follows that

$$\begin{aligned} \det(A) &= \det(A^T) \\ &= \det(-A) \\ &= (-1)^{\text{order of } A} \det(A) \\ &= -\det(A), \text{ since order of } A \text{ is odd.} \end{aligned}$$

This follows that

$$\det(A) = 0.$$

Hence we proved the Lemma.

Now Note that if we prove (b) first then automatically (a) comes true for the integer entries. Since a perfect square number is always greater equals to 0.

So let us prove (b).

Let us consider  $A$  is an  $n \times n$  real skew-symmetric matrix, and it is enough to take  $n$  is even, by the lemma.

If all the entries of the first row of  $A$  are zero, then the determinant of  $A$  is zero, and which is a perfect square. So we are done for this case.

Now assume that at least one entry of the first row of  $A$  be non-zero.

Let us consider the matrix  $A$  as

$$A = [a_{ij}]_{n \times n}.$$

To prove the aforesaid statement we will use induction on  $n$ .

For  $n = 2$  let us look at the matrix as

$$A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}, \text{ where } b \in \mathbb{Z}.$$

Then

$$\det(A) = b^2 \geq 0.$$

So for  $n = 2$  our statement is true.

For our simplicity let us assume the non-zero entry in the first row of  $A$  be  $a_{12}$ .

Now by doing column operation on  $A$  we can make the first row of  $A$  as  $[0, a_{12}, 0, 0, \dots, 0]_{1 \times n}$ .

Similarly by doing row operation we can make the first column as  $[0, -a_{12}, 0, 0, \dots, 0]_{n \times 1}$ . After making the first column as above, note that at the end of row operation the first row will be same as  $[0, a_{12}, 0, 0, \dots, 0]_{1 \times n}$ .

Then call this changed matrix as  $B$ .

Then obviously we have

$$\det(A) = \det(B).$$



Now ignore the first row and first column in  $B$ . Then we have left a  $(n-2) \times (n-2)$  matrix, say  $C$ , which is again skew-symmetric. Again notice that

$$\det(A) = \det(B) = a_{12}^2 \det(C).$$

Now by the induction hypothesis we have

$$\det(C) = R^2, \text{ where } R \in \mathbb{Z}.$$

Therefore we have

$$\det(A) a_{12}^2 R^2 = (a_{12} R)^2.$$

This follows that  $\det(A)$  is a perfect square.

This completes the proof.

Now we will prove that  $\det(A) \geq 0$ .

We know that only real eigen-value of a skew-symmetric matrix of real entries is zero.

If  $\det(A) = 0$  then we are done.

If  $\det(A) \neq 0$ . Then the eigen-values of  $A$  are all non-zero, since determinant of  $A$  is the products of all its eigen-values.

Since other than 0 there is no real eigen-value of  $A$ , in this case 0 is not possible, so all the eigen-values are of the form  $a + ib$ , where  $a, b \in \mathbb{R}$ . Since  $A$  has even order, it has even number of eigen-values.

Now  $a + ib$  is an eigen-value of  $A$  implies  $a - ib$  is also an eigen-value of  $A$ .

Then there are  $\frac{n}{2}$  pairs  $(a + ib, a - ib)$ , which are eigen values of  $A$ .

Now determinant of  $A$  is products of all these eigen values.

Therefore

$$\det(A) = \prod (a + ib)(a - ib) = \prod (a^2 + b^2) > 0.$$

This completes the proof that  $\det(A) \geq 0$  for a skew-symmetric matrix of real entries.

## Result

4 of 4

First we show that for a skew-symmetric matrix  $A$  with integer entries  $\det(A)$  is a perfect square and then proved that  $\det(A) \geq 0$  for a skew-symmetric matrix with real entries.

## Miscellaneous Problem

1. a

Let  $G$  be a group and let  $G$  acts on a set  $S$  under the operation  $*$ .

Then the orbit of an element  $x \in S$  is given by

$$G(x) = \{gx \in S : g \in G\}$$

In Sylvester's Law, the six standard matrices are as follows

First type-  $\mathbf{1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Second type-  $\mathbf{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

Third type-  $\mathbf{3} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$

Fourth type-  $\mathbf{4} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Fifth type-  $\mathbf{5} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

Sixth type-  $\mathbf{6} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Now consider the operation of the group  $GL_2$  on  $2 \times 2$  matrices by

$$P * A = P A P^t, \text{ where } A \text{ is a symmetric matrix}$$

Now let the matrix  $A$  has general form

$$A = \begin{pmatrix} x & y \\ y & z \end{pmatrix}, \text{ where } x, y, z \in \mathbb{R}$$

Then  $\det(A)$  is given by

$$\begin{aligned} \det(A) &= \begin{vmatrix} x & y \\ y & z \end{vmatrix} \\ &= xz - y^2 \end{aligned}$$

Since the determinant is a homogeneous function, so  $y^2 = xz$  is the required geometric figure.

Define new axes  $u, v$  as

$$u = \frac{x+z}{\sqrt{2}} \text{ And } v = \frac{x-z}{\sqrt{2}}$$

Then,

$$x = \frac{u+v}{\sqrt{2}} \text{ And } z = \frac{u-v}{\sqrt{2}}$$

Substitute the values of  $x, z$  in terms of  $u, v$  in the equation  $y^2 = xz$

So,

$$\begin{aligned} y^2 &= xz \\ &= \left( \frac{u+v}{\sqrt{2}} \right) \left( \frac{u-v}{\sqrt{2}} \right) \\ &= \frac{u^2 - v^2}{2} \end{aligned}$$

Hence the equation  $y^2 = xz$  transforms into  $u^2 = v^2 + 2y^2$ , which is the equation of an elliptic cone in  $\mathbb{R}^3$



Let  $P \in GL_2$  have the general form

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ where } a, b, c, d \in \mathbb{R}$$

Now corresponding to the matrix of the first type

$$\mathbf{1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Since the matrix  $A$  is congruent to  $\mathbf{1}$

Now consider the  $P\mathbf{1}P'$

$$P\mathbf{1}P' = 0$$

Thus the first type corresponds to the points  $(0, 0, 0)$  in the space

Now corresponding to the matrix of the second type

$$\mathbf{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Consider the matrix  $P\mathbf{2}P'$

$$\begin{aligned} P\mathbf{2}P' &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^2 & ac \\ ac & c^2 \end{pmatrix} \end{aligned}$$

Thus this corresponds to the point  $(a^2, ac, c^2)$  with  $ad - bc \neq 0$ .

Then  $y^2 = xz$ , with  $x, z \geq 0$  but simultaneously they both can't be zero

Now corresponding to the matrix of the third type

$$\mathbf{3} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

Consider the matrix  $P\mathbf{3}P'$

$$\begin{aligned} P\mathbf{3}P' &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -a & -c \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -a^2 & -ac \\ -ac & -c^2 \end{pmatrix} \end{aligned}$$

Thus this corresponds to the point  $(-a^2, -ac, -c^2)$  with  $ad - bc \neq 0$ .

Then  $y^2 = xz$ , with  $x, z \leq 0$  but simultaneously they both can't be zero

Now corresponding to the matrix of the fourth type

$$4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Consider the matrix  $P4P'$

$$\begin{aligned} P4P' &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} \end{aligned}$$

Thus this corresponds to the point  $(a^2 + b^2, ac + bd, c^2 + d^2)$  with  $ad - bc \neq 0$ .

Then  $xz - y^2 = (ad - bc)^2 > 0$ , with  $x, z > 0$

Now corresponding to the matrix of the fifth type

$$5 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Consider the matrix  $P5P'$

$$\begin{aligned} P5P' &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -a & -c \\ -b & -d \end{pmatrix} \end{aligned}$$

Thus this corresponds to the point  $(-a^2 - b^2, -ac - bd, -c^2 - d^2)$  with  $ad - bc \neq 0$ .

Then  $xz - y^2 = (ad - bc)^2 > 0$ , with  $x, z < 0$

Now corresponding to the matrix of the sixth type

$$6 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

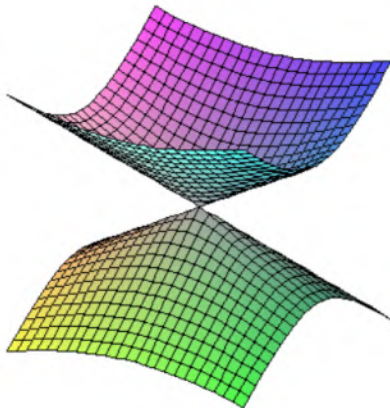
Consider the matrix  $P6P'$

$$\begin{aligned} P6P' &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ -b & -d \end{pmatrix} \\ &= \begin{pmatrix} a^2 - b^2 & ac - bd \\ ac - bd & c^2 - d^2 \end{pmatrix} \end{aligned}$$

Thus this corresponds to the point  $(a^2 - b^2, ac - bd, c^2 - d^2)$  with  $ad - bc \neq 0$ .

Then  $xz - y^2 = (ad - bc)^2 < 0$

Now, the cone is given by the following geometric figure



The cone splits into three orbits:

Type **1** corresponds to the point of origin.

Type **2** corresponds to the positive cone with  $u > 0$

Type **3** corresponds to the negative cone with  $u < 0$

Type **4** corresponds to the interior of the positive cone with  $u > 0$

Type **5** corresponds to the interior of the negative cone with  $u < 0$

Type **6** corresponds to the exterior of the cone

**Therefore, the result in the question has been proved.**

2. a

**(a)**

Since  $A$  and  $B$  symmetric then

$$A = A^t \text{ and } B = B^t.$$

Hence

$$\begin{aligned} (AB + BA)^t &= (AB)^t + (BA)^t \\ &= B^t A^t + A^t B^t \\ &= BA + AB \\ &= AB + BA, \end{aligned}$$

so that  $AB + BA$  is symmetric. Now similarly

$$\begin{aligned} (AB - BA)^t &= (AB)^t - (BA)^t \\ &= B^t A^t - A^t B^t \\ &= BA - AB \\ &= -(AB - BA), \end{aligned}$$

so that  $AB - BA$  is skew-symmetric.

**(b)**

In analogy with the **(a)** part we could show that if  $A$  and  $B$  are Hermitian then  $AB + BA$  is Hermitian and  $AB - BA$  is skew-Hermitian. However, as we're looking for symmetry, let

$$A = \begin{bmatrix} 0 & 1+i \\ 1-i & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix},$$

i.e.  $A$  and  $B$  are Hermitian. But it can be easily computed that

$$AB + BA = \begin{bmatrix} 4 & 2+2i \\ 2-2i & 4 \end{bmatrix},$$

which is neither a symmetric nor a skew-symmetric matrix. Similarly, if

$$A = \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1+i \\ 1-i & 1 \end{bmatrix},$$

then

$$AB - BA = \begin{bmatrix} 0 & -2-2i \\ 2-2i & 0 \end{bmatrix},$$

which is neither symmetric nor skew-symmetric.

(c)

We have that

$$A^t = -A \text{ and } B^t = -B,$$

so that

$$\begin{aligned}(AB + BA)^t &= (AB)^t + (BA)^t \\&= B^t A^t + A^t B^t \\&= (-B)(-A) + (-A)(-B) \\&= BA + AB \\&= AB + BA,\end{aligned}$$

so that  $AB + BA$  is symmetric. Similarly

$$\begin{aligned}(AB - BA)^t &= (AB)^t - (BA)^t \\&= B^t A^t - A^t B^t \\&= (-B)(-A) - (-A)(-B) \\&= BA - AB \\&= -(AB + BA),\end{aligned}$$

i.e.  $AB - BA$  is skew-symmetric.

(d)

We have, since  $A^t = A$  and  $B^t = -B$ ,

$$\begin{aligned}(AB + BA)^t &= (AB)^t + (BA)^t \\&= B^t A^t + A^t B^t \\&= (-B)A + A(-B) \\&= -BA - AB \\&= -(AB + BA),\end{aligned}$$

so that  $AB + BA$  is skew-symmetric. And finally

$$\begin{aligned}(AB - BA)^t &= (AB)^t - (BA)^t \\&= B^t A^t - A^t B^t \\&= (-B)A - A(-B) \\&= -BA + AB \\&= AB - BA,\end{aligned}$$

i.e.  $AB - BA$  is symmetric.

## Result

5 of 5

We obtain the results as follows:

- a.  $AB + BA$  symmetric,  $AB - BA$  skew-symmetric,
- b. Neither of  $AB + BA$  nor  $AB - BA$  symmetric or skew-symmetric (but similarly as in (a) it could be shown that the first is Hermitian and the second is skew-Hermitian),
- c.  $AB + BA$  symmetric,  $AB - BA$  skew-symmetric,
- d.  $AB + BA$  skew-symmetric,  $AB - BA$  symmetric.

3. a

## (a) Real Orthogonal

### Determinants

Let  $P$  be a real orthogonal matrix, then  $P$  is a real matrix and

$$PP^t = P^tP = I.$$

Taking a determinant of both sides of  $PP^t = I$ , and using the fact that  $\det P = \det P^t$ , we get

$$(\det P)^2 = 1,$$

which implies that the determinant of  $P$ , as it is a real number, is either 1 or  $-1$ . In order to show that these can be achieved, note that  $I$  is orthogonal and has determinant 1, while

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is orthogonal and has determinant  $-1$ .

### Eigenvalues

In the exercise **M.6**, we prove that an eigenvalue of a real orthogonal matrix must be a complex number  $\lambda$  with  $|\lambda| = 1$ . Proof that each such complex number can actually be realized as an eigenvalue of some real orthogonal matrix can be given using rotation matrices.

## (b) Unitary

### Determinants

Let  $Q$  be a unitary matrix, i.e.  $Q$  is a complex matrix and  $Q^*Q = QQ^* = I$ . Note that

$$\det Q^* = \overline{\det Q},$$

where  $\bar{\phantom{x}}$  denote the complex conjugate. This is a consequence of the fact that  $\det Q = \det Q^t$ , the additivity and multiplicativity of conjugation and the fact that a determinant of a matrix can be written as a sum of products of elements of that matrix. Now similarly as in **(a)**, by applying determinant to  $QQ^* = I$  we obtain

$$\det Q \overline{\det Q} = 1,$$

so that

$$|\det Q|^2 = 1,$$

which means that  $|\det Q| = 1$ . This means that that determinants of unitary matrices are given by complex numbers of modulus 1.

### Eigenvalues

Completely analogously as for real orthogonal matrices, it can be proved that eigenvalues of unitary matrices are complex numbers with modulus 1.



### (c) Hermitian

#### Determinants

Let  $H$  be a Hermitian matrix, so that  $H = H^*$ . Therefore, noting the equality  $\det H^* = \overline{\det H}$  for which we gave a sketch of a proof in (b), we have

$$\det H = \overline{\det H},$$

which implies that  $\det H$  is a real number. We see that each real number does indeed manifest as a determinant of a Hermitian matrix since

$$H_a = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}$$

is a Hermitian matrix for any  $a \in \mathbb{R}$ .

#### Eigenvalues

It was proved in **Theorem 8.3.11** that eigenvalues of a Hermitian matrix are real numbers.

### (d) Real symmetric, negative definite

#### Determinants

Using **Theorem 8.4.19**, which gives a characterization of positive definite symmetric matrix  $A$  according to positivity of determinant of its minors (one of which is the matrix  $A$ ), we see that if  $A$  is a real symmetric matrix then  $A$  is negative semidefinite -- as the negation of positive definite is negative semidefinite -- and  $\det A_k \leq 0$  for all minors.

In order to prove that if  $A$  is negative definite then we indeed have a sharp inequality for its determinant, i.e.  $\det A < 0$ , we just have to note that  $A$  is invertible. This follows from first observing that  $A$  is invertible iff  $-A$  is invertible, and  $-A$  is positive definite, so it is invertible by e.g. **Theorem 8.2.5**, the theorem giving the characterization of the dot product on  $\mathbb{R}^n$ .

#### Eigenvalues

First note that since Hermitian matrices have real eigenvalues, then so do symmetric matrices. Now, if  $A$  is a negative-definite symmetric matrix and  $\lambda$  is an eigenvalue of  $A$ , then there exists an  $X$  such that  $AX = \lambda X$ . By multiplying from the left by  $X^t$  we obtain

$$X^t A X = \lambda X^t X.$$

By negative definiteness, we know that the left-hand side is a negative number, and that  $X^t X > 0$ , proving that  $\lambda < 0$ ; we obtain that the eigenvalues of real symmetric, negative definite matrices must be negative real numbers.

### (e) Real skew-symmetric

#### Determinants

Let  $A$  be real and skew-symmetric  $n \times n$  matrix, so that  $A = -A^t$ . Since  $\det A = \det A^t$ , we have

$$\det A = \det -A^t = (-1)^n \det A,$$

from which it follows that if  $n$  is an odd number, then  $\det A = 0$ . If, however,  $n$  is an even number, then  $\det A \geq 0$ ; this is the statement of the **Exercise 8.4.**

#### Eigenvalues

In **Exercise 6.5**, we showed that the Spectral theorem implies that the eigenvalues of a skew-symmetric are purely imaginary, i.e. complex numbers with real part 0.



We obtain the following results:

(a), **Real orthogonal**: Determinants are either 1 or  $-1$ , eigenvalues are complex numbers  $z$  with  $|z| = 1$ .

(b), **Unitary**: Both determinants and eigenvalues are complex numbers with modulus 1.

(c), **Hermitian**: Both determinants and eigenvalues are real numbers.

(d), **Real symmetric, negative definite**: Both determinants and eigenvalues are negative real numbers.

(e), **Real skew-symmetric**: Determinants are 0 for the case of a  $n \times n$  real skew-symmetric matrix with  $n$  odd, and nonnegative real numbers if  $n$  is even, and eigenvalues are purely imaginary, i.e. complex numbers with real part 0.

[Click for more details.](#)

#### 4. a

We use a determinant identity for block matrices which states that if  $A$ ,  $B$ ,  $C$ , and  $D$  are matrices of dimension  $n \times n$ ,  $n \times m$ ,  $m \times n$ , and  $m \times m$ , and  $A$  is invertible, then

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A) \det(D - CA^{-1}B).$$

Applying this to our matrix in question, we get

$$\det \begin{bmatrix} I & E^* \\ -E & I \end{bmatrix} = \det(I) \det(I - (-E)I^{-1}(E^*)) = \det(I + EE^*).$$

Therefore, in order to show the invertibility of this matrix, which is equivalent with

$$\det \begin{bmatrix} I & E^* \\ -E & I \end{bmatrix} \neq 0$$

, it is sufficient to show that  $\det(I + EE^*) \neq 0$ , i.e. that  $I + EE^*$  is invertible. Note that  $EE^*$  is Hermitian, so we can apply the Spectral theorem for Hermitian matrices to obtain a unitary matrix  $P$  such that  $P^*(EE^*)P = D$ , where  $D$  is a diagonal matrix with diagonal entries being eigenvalues of  $EE^*$ . Now we compute

$$P^*(I + EE^*)P = P^*P + P^*(EE^*)P = I + D,$$

so that, since  $P$  and  $P^*$  are unitary and therefore invertible, we have that

$$\det(I + D) = \det(P^*(I + EE^*)P) \neq 0 \text{ if and only if } \det(I + EE^*) \neq 0.$$

It remains to prove that  $I + D$  is invertible; note that it is the diagonal matrix with entries  $1 + \lambda_i$ , where  $i$  is the  $i$ th eigenvalues of  $EE^*$ , so that

$$\det(I + D) = \prod_{i=1}^m (1 + \lambda_i),$$

i.e. it is nonzero if and only if there is no eigenvalue  $\lambda = -1$ .

In order to show that all eigenvalues of  $EE^*$  are nonnegative, let us show it is positive semidefinite. For an arbitrary (complex) vector  $X$  we have

$$X^*EE^*X = (E^*X)^*E^*X$$

which is a sum of elements of the form  $x_i \overline{x_i}$ , which is a real number greater or equal to 0, where  $x_i$  is the  $i$ th coordinate of  $E^*X$ . Therefore, by reasoning similar as that of exercise 6.14., where we showed that a symmetric matrix has positive eigenvalues iff it is positive definite, so could we show that a Hermitian matrix is positive semidefinite iff it has all its eigenvalues nonnegative, which completes our proof.

We use a determinant identity for block matrices which states that for appropriate matrices we have

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A) \det(D - CA^{-1}B),$$

after applying which we get a determinant of  $I$  plus a Hermitian matrix, which makes our problem amenable to an application of the Spectral theorem. Click to see more details.

## 5. a

Let  $a, b, c \in \mathbb{R}^3$  be an arbitrary element. Then triple vector product is defined as follows

$$(a \times b) \times c = -(c \cdot b)a + (c \cdot a)b.$$

(a)

Let  $x \in \mathbb{R}^3$ ,

Let  $T(x) = (x \times v) \times v$ , where  $v \in \mathbb{R}^3$  be fixed.

Consider  $\langle T(x), x \rangle$ ,

$$\begin{aligned} \langle T(x), x \rangle &= \langle (x \times v) \times v, x \rangle \\ &= \langle -(v \cdot v)x + (v \cdot x)v, x \rangle \\ &= \langle -\|v\|^2 x, x \rangle + \langle (v \cdot x)v, x \rangle \\ &= -\|v\|^2 \|x\|^2 + (v \cdot x)\langle v, x \rangle \end{aligned}$$

Now evaluate  $\langle x, T(x) \rangle$ ,

$$\begin{aligned} \langle x, T(x) \rangle &= \langle x, (x \times v) \times v \rangle \\ &= \langle x, -(v \cdot v)x + (v \cdot x)v \rangle \\ &= \langle x, -\|v\|^2 x \rangle + \langle x, (v \cdot x)v \rangle \\ &= -\|v\|^2 \|x\|^2 + (v \cdot x)\langle x, v \rangle \end{aligned}$$

Now clearly  $\langle x, v \rangle = \overline{\langle v, x \rangle}$ ,

But since  $x, v \in \mathbb{R}^3$  so,

$$\langle x, v \rangle = \langle v, x \rangle$$

Hence  $\langle T(x), x \rangle = \langle x, T(x) \rangle$ ,

But  $\langle T(x), x \rangle = \langle x, T^*(x) \rangle$

Use this equality so,

$$\langle x, T(x) \rangle = \langle x, T^*(x) \rangle$$

Since  $x \in \mathbb{R}^3$  is arbitrary.

So  $T = T^*$ , in particular for real space  $T = T'$ .

Thus,

**The given operator defined on a real space is symmetric.**

(b)

Let  $\beta = (e_1, e_2, e_3)$  denote the standard basis for  $\mathbb{R}^3$  such that  $e_i = (0, 0, 0, \dots, 1, 0, 0, 0)$  where 1 is at  $i^{\text{th}}$  position.

Now,

$$\begin{aligned} T(e_1) &= (e_1 \times v) \times v \\ &= ((1, 0, 0) \times (v_1, v_2, v_3)) \times (v_1, v_2, v_3) \\ &= (0, -v_3, v_2) \times (v_1, v_2, v_3) \\ &= -(v_3^2 + v_2^2), v_2v_1, v_3v_1 \end{aligned}$$

Rewrite the L.H.S with respect to the standard basis

$$\begin{aligned} T(e_1) &= -(v_3^2 + v_2^2), v_2v_1, v_3v_1 \\ &= -(v_3^2 + v_2^2)e_1 + (v_2v_1)e_2 + (v_3v_1)e_3 \end{aligned}$$

Similarly evaluate  $T(e_2)$  and  $T(e_3)$

$$T(e_2) = (v_1v_2)e_1 + -(v_3^2 + v_1^2)e_2 + (v_2v_3)e_3$$

$$T(e_3) = (v_1v_3)e_1 + (v_2v_3)e_2 + -(v_3^2 + v_1^2)e_3$$

So the matrix for the given operator is given by

$$[T]_{\beta} = \begin{pmatrix} -(v_3^2 + v_2^2) & v_1v_2 & v_1v_3 \\ v_1v_2 & -(v_3^2 + v_1^2) & v_2v_3 \\ v_1v_3 & v_2v_3 & -(v_1^2 + v_2^2) \end{pmatrix}$$

Hence, the matrix for the operator defined on the real space by  $T(x) = (x \times v) \times v$ , where

$v \in \mathbb{R}^3$  be fixed is given by

$$[T]_{\beta} = \begin{pmatrix} -(v_3^2 + v_2^2) & v_1v_2 & v_1v_3 \\ v_1v_2 & -(v_3^2 + v_1^2) & v_2v_3 \\ v_1v_3 & v_2v_3 & -(v_1^2 + v_2^2) \end{pmatrix}$$

6. a

(a)

The reasoning is correct up to the part where it goes from

$$\lambda X^t X = \lambda^{-1} X^t X$$

to  $\lambda = \lambda^{-1}$ . The assumption here is that  $X^t X \neq 0$ , which does not need to hold if  $X$  is a complex eigenvector (in fact, in the exercise 12 we are asked to prove that if  $X$  is a complex eigenvector with a complex eigenvalue then  $X^t X = 0$  **always**). To give an explicit example of a complex vector  $X$  such that  $X^t X = 0$ , take

$$X = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

(b)

As we are dealing with complex vectors, we do the natural thing and take adjoints instead of transposes. This also guarantees that  $X^*X \neq 0$  unless  $X$  is the nullvector. Now we have, in the analogy with the steps take before,

$$X^*P^*X = (PX)^*X = (\lambda X)^*X = \bar{\lambda}X^*X \quad (1)$$

and

$$X^*P^*X = X^*P^{-1}X = \lambda^{-1}X^*X \quad (2)$$

where in the second equality we used the fact that since  $P$  is real then  $P^* = P^t$  and that  $P$  is orthogonal. Now this implies that

$$\bar{\lambda} = \lambda^{-1}$$

and hence

$$\bar{\lambda}\lambda = 1.$$

This can be reformulated as  $|\lambda|^2 = 1$ , or equivalently  $|\lambda| = 1$ .

## Result

3 of 3

In (a) part we note that the error is in assuming that  $X^tX \neq 0$ , which does not need to hold if  $X$  is a complex (eigen)vector. In (b) we correct that to prove that an eigenvalue of  $P$  is a complex number with modulus 1.

7. a

!!!

8. a

Let  $P$  be any matrix such that  $P^2 = Q$ , this holds only if all the entries of the matrix  $Q$  are non-negative. In mathematical way,

$$P_{ij} = \sqrt{Q_{ij}}, \text{ where } P_{ij} \text{ and } Q_{ij} \text{ denotes the } ij^{\text{th}} \text{ element of the matrix } P \text{ and } Q.$$

(a)

Let  $A$  be a nonsingular complex matrix.

Then  $A^*$  denote the adjoint of the matrix  $A$ .

$$\text{Let } A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

Then  $A^*$  is the conjugate transpose of  $A$

So,

$$A^* = \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{m1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \cdots & \overline{a_{mn}} \end{pmatrix}$$

Now consider  $ij^{\text{th}}$  element of the matrix  $A$  and  $A^*$  which are given by

$$A_{ij} = a_{ij} \text{ and } (A^*)_{ij} = \overline{a_{ji}}$$

Hence  $A^*A$  all its entries non-negative.

So square root of the product is well defined.

$$\text{Now define } B = \sqrt{(A^*A)}$$

Then, clearly  $B^2 = A^*A$

Now consider,

$$\begin{aligned}(B^2)^* &= (A^* A)^* \\ &= A^* (A^*)^* \\ &= A^* A \\ &= B^2\end{aligned}$$

Hence from this it can be concluded that  $B$  is Hermitian & by definition of  $B$  it is positive definite.

Let  $C$  be any other matrix such that it satisfies all the above proved property.

$$\text{So, } C^2 = A^* A = B^2$$

Thus,

$$C^2 = B^2$$

Hence,

$$\Rightarrow C^2 - B^2 = O$$

Since, both  $C, B$  are positive definite.

So,

$$C = B$$

Thus,

**Let  $A$  be a nonsingular complex matrix then there exists a positive definite Hermitian matrix  $B$  such that  $B^2 = A^* A$ , moreover this matrix is uniquely determined by  $A$ .**

(b)

Let  $A$  be a nonsingular matrix.

Let  $B$  be a positive definite Hermitian matrix such that  $B^2 = A^* A$ .

Consider  $(AB^{-1})^* (AB^{-1})$

$$\begin{aligned}(AB^{-1})^* (AB^{-1}) &= (B^{-1})^* A^* AB^{-1} \\ &= (B^*)^{-1} A^* AB^{-1} \\ &= B^{-1} A^* AB^{-1}\end{aligned}$$

Now since  $B^2 = A^* A$ ,

$$\begin{aligned}(AB^{-1})^* (AB^{-1}) &= B^{-1} A^* AB^{-1} \\ &= B^{-1} B^2 B^{-1} \\ &= (B^{-1} B)(BB^{-1}) \\ &= I\end{aligned}$$

Now consider  $(AB^{-1})(AB^{-1})^*$

$$\begin{aligned}(AB^{-1})(AB^{-1})^* &= AB^{-1} (B^{-1})^* A^* \\ &= AB^{-1} (B^*)^{-1} A^* \\ &= AB^{-1} (B)^{-1} A^* \\ &= AB^{-2} A^*\end{aligned}$$

Now use  $B^2 = A^* A$

$$\begin{aligned}(AB^{-1})(AB^{-1})^* &= AB^{-2}A^* \\ &= A(B^2)^{-1}A^* \\ &= A(A^*A)^{-1}A^*\end{aligned}$$

Now,

$$\begin{aligned}(AB^{-1})(AB^{-1})^* &= A(A^*A)^{-1}A^* \\ &= AA^{-1}(A^*)^{-1}A^* \\ &= I\end{aligned}$$

Hence,

$$(AB^{-1})(AB^{-1})^* = (AB^{-1})^*(AB^{-1}) = I.$$

Thus,

$AB^{-1}$  is a unitary matrix

(c)

Let  $A$  be a nonsingular matrix of order  $m \times n$ .

Choose  $P^2 = A^* A$ , thus  $P = \sqrt{A^* A}$

Let  $\text{Rank}(A) = r \leq n$ .

By Spectral theorem, there exists an orthonormal basis of eigenvectors for  $P$  say

$$\{v_1, v_2, \dots, v_n\}.$$

This implies that  $\sqrt{A^* A}(v_i) = Pv_i = \lambda_i v_i, 1 \leq i \leq n$ , where  $\lambda_i$  denotes the eigenvalue corresponding to the eigenvector  $v_i$ .

Since,  $\text{Rank}(A) = r \leq n$

Also all the eigenvalues  $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_r > 0$  and  $\lambda_{r+1}, \lambda_{r+2}, \dots, \lambda_n = 0$

Consider the set

$$\left\{ \frac{1}{\lambda_1} Av_1, \frac{1}{\lambda_2} Av_2, \dots, \frac{1}{\lambda_r} Av_r \right\},$$

Consider the inner product of any two elements of this set,

$$\begin{aligned}\left\langle \frac{1}{\lambda_i} Av_i, \frac{1}{\lambda_j} Av_j \right\rangle &= \frac{1}{\lambda_i \lambda_j} \langle Av_i, Av_j \rangle \\ &= \frac{1}{\lambda_i \lambda_j} \langle v_i, A^* Av_j \rangle \\ &= \frac{1}{\lambda_i \lambda_j} \langle v_i, \lambda_j^2 v_j \rangle \\ \left\langle \frac{1}{\lambda_i} Av_i, \frac{1}{\lambda_j} Av_j \right\rangle &= \frac{\lambda_j}{\lambda_i} \langle v_i, v_j \rangle \\ &= 0\end{aligned}$$

Clearly the set  $\left\{ \frac{1}{\lambda_1} Av_1, \frac{1}{\lambda_2} Av_2, \dots, \frac{1}{\lambda_r} Av_r \right\}$  can be extended to include  $n-r$  more orthonormal vectors say  $\{y_{r+1}, y_{r+2}, \dots, y_n\}$ .

Define  $U$  by

$$\begin{aligned}U &= \left[ \frac{1}{\lambda_1} Av_1 \mid \frac{1}{\lambda_2} Av_2 \mid \dots \mid \frac{1}{\lambda_r} Av_r \mid y_{r+1} \mid \dots \mid y_n \right] \cdot P \\ &= \left[ \frac{1}{\lambda_1} Av_1 \mid \frac{1}{\lambda_2} Av_2 \mid \dots \mid \frac{1}{\lambda_r} Av_r \mid y_{r+1} \mid \dots \mid y_n \right] \cdot [v_1 \mid v_2 \mid \dots \mid v_n]\end{aligned}$$



Thus a  $m \times n$  matrix is obtained which has orthonormal columns.

Now let the standard unit vector  $e_i \in \mathbb{C}^m$  as

$$[e_i]_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Now consider  $Pv_i$

$$\begin{aligned} Pv_i &= [v_1 | v_2 | \dots | v_n]^* \cdot v_i \\ &= 0 + 0 + 0 + \dots + [(0, 0, \dots, 1, \dots, 0)] + 0 + \dots + 0 \\ &= e_i \end{aligned}$$

Thus,

$$Uv_i = \begin{cases} \frac{1}{\lambda_i} Av_i, & 1 \leq i \leq r \\ y_i, & r+1 \leq i \leq n \end{cases}$$

Now evaluate  $UPv_i, 1 \leq i \leq r$

$$\begin{aligned} UPv_i &= U \lambda_i v_i \\ &= \lambda_i Uv_i \\ &= \lambda_i \frac{1}{\lambda_i} Av_i \\ &= Av_i \end{aligned}$$

Also if  $r+1 \leq i \leq n$ ,

$$\begin{aligned} UPv_i &= U \lambda_i v_i \\ &= \lambda_i Uv_i \\ &= 0 \frac{1}{\lambda_i} A \\ &= 0 \end{aligned}$$

Finally  $UP = A^*$

$$\begin{aligned} A^* &= UP \\ &= U \sqrt{(A^* A)}^* \\ &= U \sqrt{AA^*} \end{aligned}$$

Thus  $A = \sqrt{AA^*} U^*$

Since  $U$  consists of orthonormal vectors its columns so it turns out to be unitary.

Hence,

**Every nonsingular matrix is a product of a positive definite Hermitian matrix and a unitary matrix.**

(d)

Let  $A = UP$ ,

Then,

$$\begin{aligned} A^* A &= (UP)^* (UP) \\ &= P^* U^* UP \\ &= P^* P \\ &= P^2 \end{aligned}$$

Thus,  $P = \sqrt{A^* A}$  implies  $P$  is unique.

Now, since  $A$  is invertible then  $U = AP^{-1}$  which implies  $U$  is also unique.

Hence

The polar decomposition of a nonsingular matrix is unique.

(e)

Let  $A \in GL_n$  be an  $n \times n$  invertible matrix.

Then  $A = U'P$ , where  $U'$  is a unitary operator and  $P$  is a positive definite Hermitian matrix.

Let  $U \in U_n$

Now consider left multiplication of  $U$  on  $A$ .

So,

$$UA = UU'P$$

But since product of two unitary operators is a unitary operator so  $UU' = Q(\text{say})$  is a unitary operator.

$$UA = QP, \text{ where } Q \text{ is a unitary group.}$$

Hence,

**It can be deduced that when  $U_n$  unitary group acts on  $GL_n$  general linear group by left multiplication then it is equivalent to a unitary group acting on a positive definite matrix.**

9. a

Let  $V$  be a Euclidean space such that  $\dim(V) = n$ . Let  $S = \{v_1, v_2, \dots, v_k\}$ .

The positive combination of the vectors of the set  $S$  is a linear combination in which scalars are positive.

Hyper plane- Let  $w(\neq 0) \in V$ , the subspace of  $V$  consisting of vectors which are orthogonal to the vector  $w$  is called a hyper plane.

Mathematically, the hyper plane is given by

$$U = \{v \in V \mid \langle v, w \rangle = 0\}$$

(a)

Let  $V$  be a Euclidean space such that  $\dim(V) = n$ . Let  $S = \{v_1, v_2, \dots, v_k\}$ .

The positive combination of the vectors of the set  $S$  is a linear combination in which scalars are positive.

Let  $w(\neq 0) \in V$ , the subspace of  $V$  consisting of vectors which are orthogonal to the vector  $w$  is called a hyper plane.

Mathematically, the hyper plane is given by

$$U = \{v \in V \mid \langle v, w \rangle = 0\}$$

Let,

$$A = S \text{ is not contained in any half-space}$$

$$B = \text{For every non-zero vector } w(\neq 0) \in V, \langle v_i, w \rangle < 0 \text{ for some } i, i = 1, 2, 3, \dots, k$$

Assume on contrary that statement  $A$  does not hold that is,  $S = \{v_1, v_2, \dots, v_k\}$  is contained in one of the half spaces.

Without loss of generality, let  $S \in \{v \in V \mid \langle v, w \rangle \geq 0\}$

Now since  $v_i \in S$  for every  $1 \leq i \leq k$

So,

$$\langle v_i, w \rangle \geq 0, \text{ for every } 1 \leq i \leq k$$

Thus  $\langle v_i, w \rangle$  is not less than zero for any  $1 \leq i \leq k$

Hence  $\sim A \Rightarrow \sim B$ , so  $B \Rightarrow A$

Now conversely assume that  $B$  does not hold that is, for any  $w(\neq 0) \in V$ ,  $\langle v_i, w \rangle \geq 0$  for every

$$1 \leq i \leq k$$

Since,  $\langle v_i, w \rangle \geq 0$  for every  $1 \leq i \leq k$

This implies that for every  $1 \leq i \leq k$

$$v_i \in \{v \in V \mid \langle v, w \rangle \geq 0\}$$

Hence  $S$  is contained in one of the half-space

Again  $\sim B \Rightarrow \sim A$ , so  $A \Rightarrow B$

**Therefore, the statements mentioned in the question are equivalent.**

**(b)**

$$\text{Let } S' = S - \{v_k\} = \{v_1, v_2, \dots, v_{k-1}\}$$

Let  $S$  does not lie in any of the half space

$$\text{Let } w(\neq 0) \in V$$

Assume that  $S'$  is contained in the half space  $\{v \in V \mid \langle v, w \rangle \geq 0\}$

Since  $S'$  lies in the half space  $\{v \in V \mid \langle v, w \rangle \geq 0\}$  so  $v_k$  can be written as positive combination of vectors of  $S'$

Mathematically,

There exist scalars  $p_1, p_2, \dots, p_{k-1}$  with  $p_i > 0 \forall 1 \leq i \leq k-1$  such that

$$v_k = \sum_{i=1}^{k-1} p_i v_i$$

Consider the inner product  $\langle v_k, w \rangle$

$$\begin{aligned} \langle v_k, w \rangle &= \left\langle \sum_{i=1}^{k-1} p_i v_i, w \right\rangle \\ &= \sum_{i=1}^{k-1} p_i \langle v_i, w \rangle \end{aligned}$$

Since  $\{v_1, v_2, \dots, v_{k-1}\} \in \{v \in V \mid \langle v, w \rangle \geq 0\}$  so  $\langle v_i, w \rangle \geq 0$ , for every  $1 \leq i \leq k$

Hence  $\langle v_k, w \rangle \geq 0$

This implies that  $v_k \in \{v \in V \mid \langle v, w \rangle \geq 0\}$ , thus  $S \in \{v \in V \mid \langle v, w \rangle \geq 0\}$  which is a contradiction the given fact.

**Therefore, the given result in the statement has been proved.**

**(c)**

Let  $S$  be not contained in any of the half-space.

Part **(b)** implies that  $\text{span}(S) = V$  and since

Let that no vector in  $S$  is a non-negative linear combination of vectors from  $S$ .

Proceed by induction on dimension of space  $V$

For  $\dim(V) = 1$

Then  $S$  has positive as well as negative scalar multiples of the only vector in space which implies that  $0$  can be made as the positive combination of  $S$ .

Assume that the result holds true for  $\dim(V) = n-1$

Now project the vectors in the set  $S' = S - \{v_k\}$  onto the space  $U = \{v \in V \mid \langle v, v_k \rangle = 0\}$

The projection map  $\pi: S' \rightarrow U$  is given as follows

$$\pi(v_i) = v_i - c_i v_k, \text{ for every } 1 \leq i \leq k-1$$

Here the scalars  $c_i$  are given by

$$c_i = \frac{\langle v_i, v_k \rangle}{\langle v_k, v_k \rangle}, \text{ for every } 1 \leq i \leq k-1$$

Then the vectors  $\{\pi(v_1), \pi(v_2), \dots, \pi(v_{k-1})\}$  belong to  $U = \{v \in V \mid \langle v, v_k \rangle = 0\}$  which is a  $n-1$  dimensional subspace of the vector space  $V$

Since no vector in  $S$  is a non-negative linear combination of vectors from  $S$ .

So, the vectors  $\{\pi(v_1), \pi(v_2), \dots, \pi(v_{k-1})\}$  do not lie in any of the half plane of  $U = \{v \in V \mid \langle v, v_k \rangle = 0\}$ .

Then apply the induction hypothesis to  $\{\pi(v_1), \pi(v_2), \dots, \pi(v_{k-1})\}$

So there exist positive scalars  $a_i$ ,  $1 \leq i \leq k-1$  such that

$$0 = \sum_{i=1}^{k-1} a_i \pi(v_i), \text{ where } a_i \geq 0 \text{ for every } 1 \leq i \leq k-1$$

Substitute the value of  $\pi(v_i)$  in above equation

$$\begin{aligned} 0 &= \sum_{i=1}^{k-1} a_i \pi(v_i) \\ &= \sum_{i=1}^{k-1} a_i (v_i - c_i v_k) \\ &= \sum_{i=1}^{k-1} a_i v_i - \left( \sum_{i=1}^{k-1} a_i c_i \right) v_k \end{aligned}$$

Now from part (a)  $\langle v_i, v_k \rangle < 0$  for every  $i$ , so  $c \leq 0$

Since  $a_i \geq 0$  and  $c \leq 0$ , so

$$-\left( \sum_{i=1}^{k-1} a_i c_i \right) \geq 0$$

Thus,

$$\begin{aligned} 0 &= \sum_{i=1}^{k-1} a_i v_i - \left( \sum_{i=1}^{k-1} a_i c_i \right) v_k \\ &= a_1 v_1 + \dots + a_{k-1} v_{k-1} - (a_1 c_1 + \dots + a_{k-1} c_{k-1}) v_k \end{aligned}$$

Now  $a_i \geq 0$  for every  $1 \leq i \leq k-1$  and  $-(a_1 c_1 + \dots + a_{k-1} c_{k-1}) \geq 0$

So  $0$  is positive combination of vectors of  $S$

Thus, (i)  $\Rightarrow$  (iii)

Now let  $\text{span}(S) = V$  and  $0$  is positive combination of vectors of  $S$

Let  $v \in V$  be arbitrary.

Then there exists scalars  $a_i$ ,  $1 \leq i \leq k$  such that

$$v = \sum_{i=1}^k a_i v_i$$

Also,  $0 = \sum_{i=1}^k b_i v_i$  where  $b_i$  are positive for every  $1 \leq i \leq k$

If  $a_i$  are not positive for every  $1 \leq i \leq k$  choose  $p$  with  $1 \leq p \leq k$  such that  $a_p$  is negative and

$$|a_p| > a_i \text{ and } a_p < a_i \text{ for every } 1 \leq i \leq k$$

Similarly choose  $q$  with  $1 \leq q \leq k$  such that

$$b_q < b_i \text{ for every } 1 \leq i \leq k$$

Then  $v \in V$  can be written as

$$\begin{aligned} v &= v - \left( \frac{a_p}{b_q} \right) 0 \\ &= \sum_{i=1}^k a_i v_i - \left( \frac{a_p}{b_q} \right) \sum_{i=1}^k b_i v_i \\ &= \sum_{i=1}^k \left( a_i - a_p \left( \frac{b_i}{b_q} \right) \right) v_i \end{aligned}$$

Now consider the coefficient of  $v_i$

Since,

$$\begin{aligned} b_q &< b_i \\ \Rightarrow 1 &< \frac{b_i}{b_q} \end{aligned}$$

Also  $a_p$  is negative, so multiplying both sides of above inequality by  $a_p$  reverses the sign

$$\begin{aligned} 1 &< \frac{b_i}{b_q} \\ \Rightarrow a_p &> \left( \frac{a_p}{b_q} \right) b_i \\ \Rightarrow -a_p &< -\left( \frac{a_p}{b_q} \right) b_i \\ \Rightarrow a_i - a_p &< a_i - \left( \frac{a_p}{b_q} \right) b_i \end{aligned}$$

Now since  $a_p < a_i$  for every  $1 \leq i \leq k$  this implies that  $a_i - a_p > 0$  for every  $1 \leq i \leq k$

Thus,

$$a_i - \left( \frac{a_p}{b_q} \right) b_i > 0, \text{ for every } 1 \leq i \leq k$$

So,

$$v = \sum_{i=1}^k \left( a_i - a_p \left( \frac{b_i}{b_q} \right) \right) v_i, \text{ where } a_i - \left( \frac{a_p}{b_q} \right) b_i > 0 \text{ for every } 1 \leq i \leq k$$

Hence  $v \in V$  is a positive combination of  $S$ .

Since  $v \in V$  was arbitrary so every vector in  $V$  is a positive combination of  $S$ .

Thus, (iii)  $\Rightarrow$  (ii)

Now let every vector in  $V$  is a positive combination of  $S$ .

Let  $w (\neq 0) \in V$

Assume on contrary that statement does not hold that is,  $S = \{v_1, v_2, \dots, v_k\}$  is contained in one of the half spaces.

Without loss of generality, let  $S \in \{v \in V \mid \langle v, w \rangle \leq 0\}$

Since otherwise  $w$  can be replaced with  $(-w)$  to make sure that  $S \in \{v \in V \mid \langle v, w \rangle \geq 0\}$

Since  $w(\neq 0) \in V$  so there exist scalars  $p_i$ ,  $1 \leq i \leq k$  such that

$$w = \sum_{i=1}^k p_i v_i, \text{ where } p_i \geq 0 \text{ for every } 1 \leq i \leq k$$

Now consider the inner product  $\langle w, w \rangle$

$$\begin{aligned} \langle w, w \rangle &= \left\langle \sum_{i=1}^k p_i v_i, w \right\rangle \\ &= \sum_{i=1}^k p_i \langle v_i, w \rangle \end{aligned}$$

Since  $S \in \{v \in V \mid \langle v, w \rangle \leq 0\}$ , so  $\langle v_i, w \rangle \leq 0$  for every  $1 \leq i \leq k$

Also  $p_i \geq 0$  for every  $1 \leq i \leq k$

This implies that  $\sum_{i=1}^k p_i \langle v_i, w \rangle \leq 0$ , that is,  $\langle w, w \rangle \leq 0$ .

But  $\langle w, w \rangle \geq 0$ , so  $\langle w, w \rangle = 0$  which implies that  $w = 0$ .

A contradiction has been attained.

This contradiction arises due to the wrong assumption that  $S = \{v_1, v_2, \dots, v_k\}$  is contained in one of the half spaces.

Hence (ii)  $\Rightarrow$  (i)

**Therefore, the statements mentioned in the question are equivalent.**

10. a

Consider the Fourier matrix denoted by  $F$  of order  $n \times n$  which runs from 0 to  $n-1$ . Let the  $ij^{th}$  entry of the matrix be given by  $\zeta^{ij}$  where  $\zeta = e^{2\pi i/n}$ .

Clearly Discrete Fourier Matrix can be written as

$$F = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & \dots & 1 & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-2} & \omega^{n-1} \\ \vdots & \omega^2 & \omega^4 & \dots & \dots & \omega^{2(n-1)} \\ \vdots & \omega^3 & \omega^6 & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \dots & \omega^{(n-1)(n-1)} \end{pmatrix}, \text{ where}$$

$\omega = e^{2\pi i/n}$  is the primitive  $n^{th}$  root of unity and  $i = \sqrt{-1}$ . The factor  $1/\sqrt{n}$  is for normalization purpose.

(a)

Consider a set of complex number  $\{b_0, b_1, \dots, b_{n-1}\}$  denoted by the matrix representation  $[b_v]$  and the set  $\{c_0, c_1, \dots, c_{n-1}\}$  denoted by the matrix representation  $[c_v]$ .

Thus to find a complex polynomial  $f(t) = c_0 + c_1 t + c_2 t^2 + \dots + c_{n-1} t^{n-1}$  such that  $f(\zeta^v) = b_v$  is equivalent to solving the following system of linear equation

$$Ac = \frac{b}{\sqrt{n}}, \text{ where } c = [c_v] \text{ and } b = [b_v].$$

Since by definition of  $A$  it is clear that  $A$  is unitary.

Hence the solution of the above system of linear equation is given as follows

$$c = \frac{1}{\sqrt{n}} A^* b.$$

Thus the Fourier matrix provides a solution to the above mentioned problem.

**Hence the Fourier matrix solves the interpolation problem: Given complex numbers**

$b_0, b_1, \dots, b_{n-1}$  find a complex polynomial  $f(t) = c_0 + c_1 t + c_2 t^2 + \dots + c_{n-1} t^{n-1}$  such that

$$f(\zeta^v) = b_v.$$



(b)

Since,

$$A = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & \dots & 1 & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-2} & \omega^{n-1} \\ \vdots & \omega^2 & \omega^4 & \dots & \dots & \omega^{2(n-1)} \\ \vdots & \omega^3 & \omega^6 & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \dots & \omega^{(n-1)(n-1)} \end{pmatrix}$$

Hence,  $A' = A$

Thus matrix  $A$  is symmetric.

Now,

$$A^* A = A^* A.$$

Since  $A' = A$

This implies that  $A$  is normal.

So the  $ij^{th}$  element of the product  $A^* A$  is same as the  $ij^{th}$  element of the  $A^2$ .

Thus,

$$\begin{aligned} A^2_{ij} &= \frac{1}{n} \sum_k \zeta^{k(i+j)} \\ &= \begin{cases} 1, i+j \equiv 0 \pmod{n} \\ 0, \text{otherwise} \end{cases} \end{aligned}$$

Hence,

$A$  is a symmetric and is a normal matrix. And  $A^2$  is given by

$$\begin{aligned} A^2_{ij} &= \frac{1}{n} \sum_k \zeta^{k(i+j)} \\ &= \begin{cases} 1, i+j \equiv 0 \pmod{n} \\ 0, \text{otherwise} \end{cases} \end{aligned}$$

(c)

Now use the result of the part (b) that is,

$$\begin{aligned} A^2_{ij} &= \frac{1}{n} \sum_k \zeta^{k(i+j)} \\ &= \begin{cases} 1, i+j \equiv 0 \pmod{n} \\ 0, \text{otherwise} \end{cases} \end{aligned}$$

Evaluate  $A^4$

$$(A^2)^2 = \begin{cases} 1, i=j \\ 0, \text{otherwise} \end{cases}$$

Thus,  $A^4 = I$

Use this equation to determine the eigenvalues of  $A$ .

Let  $\lambda_i$  be an arbitrary eigenvalue of  $A$ .

Since,  $A^4 = I$

So,

$$\lambda^4 = 1$$

The possible solutions of this equation are  $1, -1, i, -i$  where  $i = \sqrt{-1}$

Thus the set of eigenvalues for the above defined Fourier matrix are  $\{1, -1, i, -i\}$ .

11. a

To prove that  $A$  defines an orthogonal projection to its image  $W$  if and only if  $A^2 = A$   
 $= A' A$

[Comment](#)

Step 2 of 4 ^

Suppose that an orthogonal projection map  $P$  to  $V$ , where  $V$  is vector space.

Suppose that  $A$  defines an orthogonal projection to its image  $W$ , where  $A$  is symmetric.

Define a unique map  $x = v + w \mapsto x'$  for  $x \in V$  with unique decomposition  $v$  and  $w$ .

$$\begin{aligned} A^2 x &= A(Ax) \\ &= Ax' \\ &= x' \\ &= Ax \end{aligned}$$

This implies that,  $A^2 = A$

$$\begin{aligned} A^2 x &= A(Ax) \\ A^2 x &= A'(Ax) \\ A^2 &= A' A \end{aligned}$$

Hence,

$$\begin{aligned} A^2 &= A \\ &= A' A \end{aligned}$$

**Conversely,**

Suppose,

$$\begin{aligned} A^2 &= A \\ &= A' A \end{aligned}$$

Let  $x \in V$ , then

$$v = Av + (I - A)v$$

Where,  $Av \in \text{im } A$  and  $(I - A)x \in \ker A$

Now, take  $x \in \ker A \cap \text{im } A$

$$x \in \text{im } A$$

Then there exist  $x'$  such that  $Ax' = x$

And,

$$x \in \ker A$$

This implies that,

$$Ax = 0$$

This implies that,

$$\begin{aligned} AAx' &= 0 \\ Ax' &= 0 \\ x &= 0 \end{aligned}$$

This implies that,  $\ker A \cap \text{im } A = \{0\}$

This implies that,  $V = (\ker A) \oplus (\text{im } A)$  but  $\ker A$  and  $\text{im } A$  are orthogonal.

**Hence,**  $A$  is an orthogonal projection to its image  $W$ .

12. a

!!!

13. a

!!!

14. a

Let us first prove that we cannot have  $n + 2$  such points. Suppose to the contrary, that we have  $n + 2$  vectors  $v_1, v_2, \dots, v_{n+2}$  such that  $v_i \cdot v_j < 0$  for all  $i \neq j$ . Note that, since  $\mathbb{R}^n$  is  $n$ -dimensional, we can pick  $n + 2$  scalars,  $c_1, \dots, c_{n+2}$ , such that

$$\sum_{i=1}^{n+2} c_i v_i = 0, \quad (1)$$

but also that some  $c_i > 0$  and some  $c_i < 0$ . Note that this part of the argument does not go through for  $n + 1$  points, as since  $v_1, \dots, v_n$  might be linearly independent, then there are unique  $c'_1, \dots, c'_n$  such that  $\sum_{i=1}^n c'_i v_i = v_{n+1}$ , and hence they might all be negative or positive. Therefore, let  $I$  be the set of indices such that  $c_i > 0$  if  $i \in I$  and  $I'$  the set of indices such that  $c_i < 0$  if  $i \in I'$ , so that we can rewrite (1) as

$$\sum_{i \in I} c_i v_i = \sum_{i' \in I'} d_{i'} v_{i'}, \quad (2)$$

with  $d_{i'} > 0$ . By the hypothesis, since we have  $v_i \cdot v_{i'} < 0$  for any  $i \in I$  and  $i' \in I'$ , we have that the dot product of the left-hand side and the right-hand side of (2) is negative; but note that it is also, as we're taking the dot product of a vector with itself, a square of the length of that vector (i.e. the left-hand side and the right-hand side). But this is a contradiction, since the square of the length cannot be negative, hence proving that the maximum number is strictly smaller than  $n + 2$ .

The proof of existence relies on existence of regular  $n$ -dimensional simplices centered on origin. They can be constructed by noting its two properties: the distances of its vertices (of which there are  $n + 1$ ) to its center are equal, and the angle subtended by any pair of its vertices through its center is  $\arccos\left(\frac{-1}{n}\right)$ . This second property implies that the dot product between any pair of its vertices (seen as vectors) is  $-1/n$ , which is what we needed.

## Result

3 of 3

We first prove that such number is strictly smaller than  $n + 2$ , and then note that vertices of the regular  $n$ -dimensional simplex centered on origin give such a collection of  $n + 1$  vectors. Click to see more details.

15. a

!!!

# 9

## Chapter 9

### Section 1

1. a

Consider the groups  $GL_n(\mathbb{C})$  and  $GL_{2n}(\mathbb{R})$ .

(a)

Yes,  $GL_n(\mathbb{C})$  is isomorphic to a subgroup of  $GL_{2n}(\mathbb{R})$ .

Since,

The orthogonal group  $O(2n, \mathbb{R})$  is a subgroup of  $GL_{2n}(\mathbb{R})$ ,

As a notation,

$$O(2n, \mathbb{R}) < GL_{2n}(\mathbb{R}).$$

Now,

Assume,

$$\varphi: GL_n(\mathbb{C}) \rightarrow O(2n, \mathbb{R})$$

Since,

$$A \in O(2n, \mathbb{R})$$

Where,

$$A^T A = I$$

Now,

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}^T = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

As a conjugate to the element  $a + ib \in \mathbb{C}$ ,

Therefore,

$$A^T = \overline{A}.$$

Hence,  $\boxed{GL_n(\mathbb{C}) \cong O(2n, \mathbb{R}) < GL_{2n}(\mathbb{R})}$ .

(b)

Yes, the group  $SO_2(\mathbb{C})$  is a bounded subset of  $\mathbb{C}^{2 \times 2}$ .

Assume,

$$\begin{bmatrix} a+ib & -(c+id) \\ c+id & a+ib \end{bmatrix} \in SO_2(\mathbb{C})$$

And,

$$\begin{bmatrix} a+ib & -(c+id) \\ c+id & a+ib \end{bmatrix} \begin{bmatrix} a+ib & -(c+id) \\ c+id & a+ib \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then,

$$\begin{bmatrix} a^2 - b^2 + c^2 - d^2 + 2i(ab + cd) & 0 \\ 0 & a^2 - b^2 + c^2 - d^2 + 2i(ab + cd) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So,

$$\begin{aligned} a^2 - b^2 + c^2 - d^2 &= 1 \\ ab + cd &= 0 \end{aligned}$$

Also,

$$\det \begin{bmatrix} a+ib & -(c+id) \\ c+id & a+ib \end{bmatrix} = 1$$

Therefore,

$$\begin{aligned} a^2 - b^2 + c^2 - d^2 + 2(ab + cd) &= 1 \\ a^2 + b^2 + 2ab + c^2 + d^2 + 2cd &= 1 + 2(b^2 + d^2) \\ (a+b)^2 + (c+d)^2 &= 1 + 2(b^2 + d^2) \end{aligned}$$

As,  $a, b, c$  and  $d$  are bounded by above relation.

Hence, the group  $SO_2(\mathbb{C})$  is a bounded subset of  $\mathbb{C}^{2 \times 2}$ .

2. a

Consider the properties of the columns of a matrix in the Lorentz group  $O_{3,1}$ .

The Lorentz group,

$$O_{3,1} = \{P \in GL_n : P^T I_{3,1} P = I_{3,1}\}$$

Assume,

$$P = \begin{bmatrix} X & Y & Z & W \end{bmatrix}$$

Then,

$$P^T = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

So,

$$\begin{aligned} X^2 &= 1, \\ Y^2 &= 1, \\ Z^2 &= 1, \\ W^2 &= 1. \end{aligned}$$

And,

The remaining column products are zero.

Therefore, the columns of a matrix in the Lorentz group also form an orthonormal basis.

3. a

Consider the orthogonal group  $O_4$  and the Lorentz group  $O_{3,1}$ .

[Comment](#)

Step 2 of 2 ^

The Lorentz form,

$$= Y^T I_{3,1} Y$$

$$= y_1^2 + y_2^2 + y_3^2 - y_4^2$$

It is isomorphic to  $S^2 \times S^1$ .

The orthogonal form,

$$= Y^T Y$$

$$= y_1^2 + y_2^2 + y_3^2 + y_4^2$$

It is isomorphic to  $S^3$ .

So,

Both, the Lorentz form and the orthogonal form will be same if  $y_4 = 0$ .

Since,

There is no continuous isomorphism between  $S^3$  and  $S^2 \times S^1$ .

Therefore, **there is no continuous isomorphism from the orthogonal group  $O_4$  to the Lorentz group  $O_{3,1}$ .**

4. a

Consider the group  $O_{1,1}$ .

Let  $P \in O_{1,1}$  such that,

$$P = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then,

By definition of  $O_{1,1}$ ,

$$\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

After matrix multiplication,

$$\begin{bmatrix} a_{11} & -a_{21} \\ a_{12} & -a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Now,

$$\begin{bmatrix} a_{11}^2 - a_{21}^2 & a_{11}a_{12} - a_{21}a_{22} \\ a_{11}a_{12} - a_{21}a_{22} & a_{12}^2 - a_{22}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Thus,

$$a_{11}^2 - a_{21}^2 = 1$$

$$a_{12}^2 - a_{22}^2 = -1$$

$$a_{11}a_{12} - a_{21}a_{22} = 0$$

Assume  $a_{11}$  an arbitrary constant  $c$ ,

$$a_{11} = c$$



As a result, four components are:

$$\begin{bmatrix} -c & \sqrt{c^2-1} \\ \sqrt{c^2-1} & c \end{bmatrix}, \begin{bmatrix} c & -\sqrt{c^2-1} \\ \sqrt{c^2-1} & c \end{bmatrix}, \begin{bmatrix} c & \sqrt{c^2-1} \\ -\sqrt{c^2-1} & c \end{bmatrix}, \begin{bmatrix} c & \sqrt{c^2-1} \\ \sqrt{c^2-1} & -c \end{bmatrix}$$

As,

These four components may be mapped to one of the longitudes of  $S^3$ .

Therefore,

Any two components can be joined by a continuous path entirely in the group.

Hence, **the group  $O_{1,1}$  has four path connected components.**

5. a

Consider,

$$P \in SP_2.$$

And,

$$P = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Then,

$$P'RP = R.$$

Where,

$$R = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}.$$

Since,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} & 1 \\ -1 & \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$$

Then,

$$\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} & 1 \\ -1 & \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$$

Now,

$$\begin{bmatrix} -a_{21} & a_{11} \\ -a_{22} & a_{12} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$$

Thus,

$$\begin{bmatrix} a_{11}a_{22} - a_{21}a_{12} \\ a_{12}a_{21} - a_{11}a_{22} \end{bmatrix} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$$

So,

$$a_{11}a_{22} - a_{21}a_{12} = 1$$

Also,

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = 1.$$

As,

Both the groups  $SP_2$  and  $SL_2$  have same type of elements.

Therefore,

$$SP_2 = SL_2.$$

Assume,

$$Q = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}.$$

Then,

$$\det(Q) = 1.$$

So,

$$Q \in SL_4.$$

Now,

$$\begin{aligned} & \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \\ &= \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \\ &= \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix} \end{aligned}$$

As,

$$\begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \neq \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix}$$

Thus,

$$Q \notin SP_4.$$

Therefore,

$$SP_4 \neq SL_4.$$

Hence, **by both the results**  $SP_2 = SL_2$  **but**  $SP_4 \neq SL_4$ .

6. a

Consider the matrices,

$$\begin{bmatrix} & -I \\ I & \end{bmatrix}, \begin{bmatrix} I & B \\ & I \end{bmatrix}, \begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix}.$$

[Comment](#)

Step 2 of 4 ^

Since,

$$P = \begin{bmatrix} & -I \\ I & \end{bmatrix}$$

And,

$$S = \begin{bmatrix} & I \\ -I & \end{bmatrix}$$

Then,

$$\begin{aligned} P'SP &= \begin{bmatrix} & I \\ -I & \end{bmatrix} \begin{bmatrix} I & \\ & -I \end{bmatrix} \begin{bmatrix} & -I \\ I & \end{bmatrix} \\ &= \begin{bmatrix} & I \\ -I & \end{bmatrix} \begin{bmatrix} I & \\ & -I \end{bmatrix} \\ &= \begin{bmatrix} & I \\ -I & \end{bmatrix} \\ &= S \end{aligned}$$

Therefore, the matrix  $\begin{bmatrix} & -I \\ I & \end{bmatrix}$  is symplectic.

Since,

$$P = \begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix}$$

And,

The matrix  $A$  is invertible.

Also,

$$S = \begin{bmatrix} & I \\ -I & \end{bmatrix}$$

Then,

$$\begin{aligned} \begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix} \begin{bmatrix} I & \\ & -I \end{bmatrix} \begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix} &= \begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix} \begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix} \\ &= \begin{bmatrix} I & \\ & -I \end{bmatrix} \\ &= S \end{aligned}$$

Therefore, the matrix  $\begin{bmatrix} A' & \\ & A^{-t} \end{bmatrix}$  is symplectic.

Since,

$$P = \begin{bmatrix} I & B \\ & I \end{bmatrix}$$

And,

$$B = B'$$

Also,

$$S = \begin{bmatrix} & I \\ -I & \end{bmatrix}$$

Then,

$$\begin{aligned} \begin{bmatrix} I & \\ B & I \end{bmatrix} \begin{bmatrix} & I \\ -I & \end{bmatrix} \begin{bmatrix} I & B \\ & I \end{bmatrix} &= \begin{bmatrix} & I \\ -I & B \end{bmatrix} \begin{bmatrix} I & B \\ & I \end{bmatrix} \\ &= \begin{bmatrix} & I \\ -I & \end{bmatrix} \\ &= S \end{aligned}$$

Therefore, the matrix  $\begin{bmatrix} I & B \\ & I \end{bmatrix}$  is symplectic.

7. a

**Solution:** We will now derive the formula for the inverse of the Stereographic projection  $\pi : \mathbb{S}^3 \rightarrow \mathbb{R}^3$ . Now we have

$$\mathbb{S}^3 := \{(x, y, w, z) \mid x^2 + y^2 + w^2 + z^2 = 1 \text{ and } x, y, w, z \in \mathbb{R}\}.$$

Now the map  $\pi$  is defined by the assignment

$$\pi(x, y, w, z) = \left( \frac{x}{1-z}, \frac{y}{1-z}, \frac{w}{1-z} \right), \quad (x, y, w, z) \in \mathbb{S}^3.$$

Now for our simplicity consider

$$X = \frac{x}{1-z}, \quad Y = \frac{y}{1-z}, \quad W = \frac{w}{1-z}.$$

Then notice that

$$\begin{aligned} X^2 + Y^2 + W^2 &= \left( \frac{x}{1-z} \right)^2 + \left( \frac{y}{1-z} \right)^2 + \left( \frac{w}{1-z} \right)^2 \\ &= \frac{x^2 + y^2 + w^2}{(1-z)^2} \\ &= \frac{1-z^2}{(1-z)^2}, \quad \text{since } (x, y, w, z) \in \mathbb{S}^3 \\ &= \frac{1+z}{1-z}. \end{aligned}$$

Therefore

$$\begin{aligned}\frac{1+z}{1-z} &= X^2 + Y^2 + W^2 \\ \Rightarrow z &= \frac{X^2 + Y^2 + W^2 - 1}{X^2 + Y^2 + W^2 + 1}.\end{aligned}$$

Since  $z = \frac{X^2 + Y^2 + W^2 - 1}{X^2 + Y^2 + W^2 + 1}$  notice that

$$\begin{aligned}X &= \frac{x}{1-z} \Rightarrow x = (1-z)X = \frac{2X}{X^2 + Y^2 + W^2 + 1} \\ Y &= \frac{y}{1-z} \Rightarrow y = (1-z)Y = \frac{2Y}{X^2 + Y^2 + W^2 + 1} \\ \text{and } W &= \frac{w}{1-z} \Rightarrow w = (1-z)W = \frac{2W}{X^2 + Y^2 + W^2 + 1}.\end{aligned}$$

So the transformation coordinates are given by

$$(x, y, w, z) = \left( \frac{2X}{X^2 + Y^2 + W^2 + 1}, \frac{2Y}{X^2 + Y^2 + W^2 + 1}, \frac{2W}{X^2 + Y^2 + W^2 + 1}, \frac{X^2 + Y^2 + W^2 - 1}{X^2 + Y^2 + W^2 + 1} \right).$$

Now recall from the map  $\pi$

$$(x, y, w, z) = \pi^{-1}(X, Y, W).$$

Therefore we have

$$\pi^{-1}(X, Y, W) = \left( \frac{2X}{X^2 + Y^2 + W^2 + 1}, \frac{2Y}{X^2 + Y^2 + W^2 + 1}, \frac{2W}{X^2 + Y^2 + W^2 + 1}, \frac{X^2 + Y^2 + W^2 - 1}{X^2 + Y^2 + W^2 + 1} \right).$$

So for any  $(x, y, w) \in \mathbb{R}^3$  the inverse of the map  $\pi$ , ( Stereographic Projection ) is given by

$$\pi^{-1}(x, y, w) = \left( \frac{2x}{x^2 + y^2 + w^2 + 1}, \frac{2y}{x^2 + y^2 + w^2 + 1}, \frac{2w}{x^2 + y^2 + w^2 + 1}, \frac{x^2 + y^2 + w^2 - 1}{x^2 + y^2 + w^2 + 1} \right).$$

This completes the formula.

## Result

3 of 3

Considering  $(x, y, z, w) \in \mathbb{S}^3$  we have derive the formula for the inverse of Stereographic Projection.

## Section 2

### 1. a

Consider the formula for the inverse of the stereographic projection  $\pi: \mathbb{S}^3 \rightarrow \mathbb{R}^3$ .

Since,

The formula for stereographic projection  $\pi : S^3 \rightarrow \mathbb{R}^3$ :

$$\pi(x) = (v_1, v_2) = \left( \frac{x_1}{1-x_0}, \frac{x_2}{1-x_0} \right).$$

So,

$$v_1 = \frac{x_1}{1-x_0}, v_2 = \frac{x_2}{1-x_0}.$$

After solving for  $x_0, x_1$  and  $x_2$ .

$$x_0 = \frac{v_1^2 + v_2^2 - 1}{v_1^2 + v_2^2 + 1}, x_1 = \frac{2v_1}{v_1^2 + v_2^2 + 1}, x_2 = \frac{2v_2}{v_1^2 + v_2^2 + 1}.$$

Therefore,

The formula for the inverse of the stereographic projection  $\pi^{-1} : \mathbb{R}^3 \rightarrow S^3$ :

$$\pi^{-1}(x) = (x_0, x_1, x_2) = \left( \frac{v_1^2 + v_2^2 - 1}{v_1^2 + v_2^2 + 1}, \frac{2v_1}{v_1^2 + v_2^2 + 1}, \frac{2v_2}{v_1^2 + v_2^2 + 1} \right).$$

Hence, **the formula for the inverse of the stereographic projection  $\pi^{-1} : \mathbb{R}^3 \rightarrow S^3$  is**

$$\pi^{-1}(x) = \left( \frac{v_1^2 + v_2^2 - 1}{v_1^2 + v_2^2 + 1}, \frac{2v_1}{v_1^2 + v_2^2 + 1}, \frac{2v_2}{v_1^2 + v_2^2 + 1} \right).$$

## 2. a

Consider the parametrization for the proper subspaces of  $\mathbb{R}^2$  by a circle in two ways.

First, a subspace  $W$  intersects the horizontal axis with angle  $\theta$  and use the double angle  $\alpha = 2\theta$ .

Second, a nonzero vector  $(y_1, y_2)$  in  $W$  and consider inverse of stereographic projection to map

the slope  $\lambda = \frac{y_2}{y_1}$  to a point of  $S^1$ .

[Comment](#)

### Step 2 of 2 ^

Since,

The inverse of stereographic projection is the identity map. It maps the region  $\left\{ \frac{y_2}{y_1} \leq 1 \right\}$  in  $W$  to the southern hemisphere bijectively.

And,

The northern hemisphere to the region  $\left\{ \frac{y_2}{y_1} > 1 \right\}$  except that the north-pole is missing from this.

So,

This provides a second way to build the sphere topologically, as the union of two regions glued together.

Hence, **the second way provides a topological way to build the sphere, as the union of two regions.**

## 3. a

Stereographic Projection- Let  $S^n$  be a  $n$ -sphere in  $\mathbb{R}^{n+1}$  and let  $V$  be a  $n$ -space. Choose a North Pole say  $N = (0, 0, \dots, 1)$

Define a map  $\pi : S^n \rightarrow V$  given by

$$\pi(x) = \left( \frac{x_1}{1-x_0}, \frac{x_2}{1-x_0}, \dots, \frac{x_n}{1-x_0} \right), \text{ where } x = (x_0, x_1, \dots, x_n) \in S^n$$



(a)

Let  $z = z_1 + z_2 i$  be a point in the complex plane.

Consider the line joining North Pole  $N$  and  $z = z_1 + z_2 i$

The equation of such line is given by

$$(x_1, x_2, x_3) = (0, 0, 1) + s(z_1, z_2, -1)$$

Now the point of the line which lies on the sphere  $S^2$  is a point for which  $x_1^2 + x_2^2 + x_3^2 = 1$  holds

Use above equation to determine the values of  $x_1$ ,  $x_2$  and  $x_3$  which lie on the sphere  $S^2$

$$\begin{aligned}x_1^2 + x_2^2 + x_3^2 &= 1 \\x_1 &= tz_1 \\x_2 &= tz_2 \\x_3 &= 1 - s\end{aligned}$$

Then,

$$\begin{aligned}x_1^2 + x_2^2 + x_3^2 &= 1 \\(sz_1)^2 + (sz_2)^2 + (1-s)^2 &= 1 \\s^2(z_1^2 + z_2^2) + (1-s)^2 &= 1\end{aligned}$$

The possible solution of this equation are given by

$$s = 0 \text{ And } s = \frac{2}{(z_1^2 + z_2^2 + 1)}$$

Now  $s = 0$  corresponds to the North Pole.

$$\text{So choose the other possibility } s = \frac{2}{(z_1^2 + z_2^2 + 1)} = \frac{2}{(|z|^2 + 1)}$$

Now since  $x_1 = tz_1$ ,  $x_2 = tz_2$  and  $x_3 = 1 - s$  so

$$x_1 = \frac{2z_1}{(|z|^2 + 1)}, \quad x_2 = \frac{2z_2}{(|z|^2 + 1)} \text{ and } x_3 = \frac{(|z|^2 - 1)}{(|z|^2 + 1)} \text{ where } z_1 = \operatorname{Re} z \text{ and } z_2 = \operatorname{Im} z. \text{ So, the}$$

inverse of the stereographic projection which is denoted by say  $\sigma$  is given by

$$\sigma(z) = (x_1, x_2, x_3) \text{ Where } x_1 = \frac{2 \operatorname{Re} z}{(|z|^2 + 1)}, \quad x_2 = \frac{2 \operatorname{Im} z}{(|z|^2 + 1)} \text{ and } x_3 = \frac{(|z|^2 - 1)}{(|z|^2 + 1)}.$$

Now consider the expression  $2/(1 - x_3)$

$$2/(1 - x_3) = \frac{2}{1 - x_3}$$

Now substitute the value of  $x_3$  in the above expression

$$\begin{aligned}2/(1 - x_3) &= \frac{2}{1 - x_3} \\&= \frac{2}{1 - \left( \frac{(|z|^2 - 1)}{(|z|^2 + 1)} \right)} \\&= \frac{2(|z|^2 + 1)}{2} \\&= |z|^2 + 1\end{aligned}$$

Consider  $x_1 + ix_2$

$$\begin{aligned} x_1 + ix_2 &= \frac{2z_1}{(|z|^2 + 1)} + i \frac{2z_2}{(|z|^2 + 1)} \\ &= \frac{2 \operatorname{Re} z}{(|z|^2 + 1)} + \frac{2 \operatorname{Im} z}{(|z|^2 + 1)} \\ &= \frac{2(\operatorname{Re} z + i \operatorname{Im} z)}{(|z|^2 + 1)} \\ &= \frac{2z}{|z|^2 + 1} \end{aligned}$$

Now substitute the evaluated value of  $|z|^2 + 1$  in above derived expression so

$$\begin{aligned} x_1 + ix_2 &= \frac{2z}{|z|^2 + 1} \\ &= \frac{2z}{2/(1 - x_3)} \\ &= z(1 - x_3) \end{aligned}$$

Thus the value of  $z$  is given by

$$z = \frac{x_1 + ix_2}{1 - x_3}$$

Therefore, for  $z = z_1 + z_2 i$  stereographic projection of a point  $x = (x_1, x_2, x_3) \in S^3$  is given

$$\pi(x) = z \text{ where } z = \frac{x_1 + ix_2}{1 - x_3} \text{ and the inverse of the stereographic projection is}$$

given by  $\sigma(z)$  defined by  $\sigma(z) = (x_1, x_2, x_3)$  where  $x_1 = \frac{2 \operatorname{Re} z}{(|z|^2 + 1)}$ ,  $x_2 = \frac{2 \operatorname{Im} z}{(|z|^2 + 1)}$  and

$$x_3 = \frac{(|z|^2 - 1)}{(|z|^2 + 1)}.$$

(b)

Let  $\lambda = y_2/y_1$  where the vector  $(y_1, y_2)$  is a unit vector

As evaluated above in part (a) the map  $\sigma(z)$  is given by

$$\sigma(z) = (x_1, x_2, x_3), \text{ where } x_1 = \frac{2 \operatorname{Re} z}{(|z|^2 + 1)}, x_2 = \frac{2 \operatorname{Im} z}{(|z|^2 + 1)} \text{ and } x_3 = \frac{(|z|^2 - 1)}{(|z|^2 + 1)}$$

So  $\sigma(\lambda)$  is given by

$$\sigma(\lambda) = (x_1, x_2, x_3), \text{ where } x_1, x_2, x_3 \text{ are defined by}$$

$$x_1 = \frac{2 \operatorname{Re} \lambda}{(|\lambda|^2 + 1)}$$

$$x_2 = \frac{2 \operatorname{Im} \lambda}{(|\lambda|^2 + 1)}$$

$$x_3 = \frac{(|\lambda|^2 - 1)}{(|\lambda|^2 + 1)}$$

Since  $\lambda = y_2/y_1$  so

$$\begin{aligned} x_1 &= \frac{2 \operatorname{Re} \lambda}{(|\lambda|^2 + 1)} \\ &= \frac{2 \operatorname{Re}(y_2/y_1)}{(|y_2/y_1|^2 + 1)} \\ &= \frac{2 \operatorname{Re}(y_2/y_1)}{(|y_2|^2 + |y_1|^2)} \\ &\quad |y_1|^2 \\ &= 2|y_1|^2 \operatorname{Re}(y_2/y_1) \end{aligned}$$

On similar calculations  $x_2 = 2|y_1|^2 \operatorname{Im}(y_2/y_1)$  and  $x_3 = |y_2|^2 - |y_1|^2$

**Therefore, for unit vector  $(y_1, y_2)$ ,  $\sigma(\lambda)$  where  $\lambda = y_2/y_1$  is given by**

$$\sigma(\lambda) = (2|y_1|^2 \operatorname{Re}(y_2/y_1), 2|y_1|^2 \operatorname{Im}(y_2/y_1), |y_2|^2 - |y_1|^2) .$$

(c)

Let  $(x_1, x_2, x_3)$  correspond to the point  $(y_1, y_2)$  in subspace  $W$  and let  $(a_1, a_2, a_3)$  correspond to the point  $(y_1^*, y_2^*)$  in subspace  $W'$ .

Let  $(y_1, y_2)$  and  $(y_1^*, y_2^*)$  be orthogonal.

From part (a)

$$y_1 = \frac{x_1}{1-x_3} \text{ And } y_2 = \frac{x_2}{1-x_3}$$

Similarly from part (a)

$$y_1^* = \frac{a_1}{1-a_3} \text{ And } y_2^* = \frac{a_2}{1-a_3}$$

Since  $(y_1, y_2)$  and  $(y_1^*, y_2^*)$  are orthogonal with respect to the standard Hermitian form in  $\mathbb{C}^2$  so

$$\begin{aligned} y_1^* \overline{y_2} + y_2^* \overline{y_1} &= 0 \\ \left( \frac{a_1}{1-a_3} \right) \overline{\left( \frac{x_2}{1-x_3} \right)} + \left( \frac{a_2}{1-a_3} \right) \overline{\left( \frac{x_1}{1-x_3} \right)} &= 0 \\ \frac{1}{(1-a_3)(1-\overline{x_3})} [a_1 \overline{x_2} + a_2 \overline{x_1}] &= 0 \\ a_1 \overline{x_2} + a_2 \overline{x_1} &= 0 \end{aligned}$$

Hence the first and the second co-ordinate of  $(x_1, x_2, x_3)$  and  $(a_1, a_2, a_3)$  are orthogonal with respect to the standard Hermitian form in  $\mathbb{C}^2$ .

**Therefore, for orthogonal points in the subspace  $W$  and  $W'$  the corresponding points in  $S^2$  are te set of all those points whose first and the second co-ordinate are orthogonal to each other with respect to standard Hermitian form in  $\mathbb{C}^2$ .**

## Section 3

1. a

Consider  $P$  and  $Q$  be elements of  $SU_2$ , represented by the real vectors  $(x_0, x_1, x_2, x_3)$  and  $(y_0, y_1, y_2, y_3)$ , respectively.

Since,

$$P = \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix}$$

And,

$$Q = \begin{bmatrix} y_0 + y_1 i & y_2 + y_3 i \\ -y_2 + y_3 i & y_0 - y_1 i \end{bmatrix}$$

So,

$$\begin{aligned} PQ &= \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix} \begin{bmatrix} y_0 + y_1 i & y_2 + y_3 i \\ -y_2 + y_3 i & y_0 - y_1 i \end{bmatrix} \\ &= \begin{bmatrix} z_0 + z_1 i & z_2 + z_3 i \\ -z_2 + z_3 i & z_0 - z_1 i \end{bmatrix} \end{aligned}$$

Where,

$$z_0 = x_0 y_0 - x_1 y_1 - x_2 y_2 - x_3 y_3,$$

$$z_1 = x_1 y_0 + x_0 y_1 - x_3 y_2 + x_2 y_3,$$

$$z_2 = x_0 y_2 - x_1 y_3 + x_2 y_0 + x_3 y_1,$$

$$z_3 = x_0 y_3 + x_1 y_2 - x_2 y_1 + x_3 y_0.$$

Therefore, the real vector that corresponds to the product  $PQ$  is  $(z_0, z_1, z_2, z_3)$ .

## 2. a

Consider the group  $U_2$ .

[Comment](#)

Step 2 of 2 ^

Since,

There is a bijective correspondence between the group  $SU_2$  and the unit 3-sphere.

And,

The only spheres on which the continuous group laws can be defined are that 1-sphere and the 3-sphere.

Also,

The group  $U_2$  has two more variables than the group  $SU_2$ .

Therefore,

These two variables can be related to 1-sphere.

Hence, the group  $U_2$  is homeomorphic to the product  $S^3 \times S^1$ .

## 3. a

Consider the group  $SU_2$ .

[Comment](#)

Step 2 of 2 ^

As analogues of the longitude curves, there are great circles through the north-pole.

The great circles are the intersections of the 3-sphere with two-dimensional subspaces  $W$  of  $\mathbb{R}^4$  that contain the pole.

The intersection  $L = W \cap S^3$  will be the unit circle in  $W$ , call  $L$  a longitude.

Every element of  $SU_2$  except  $\pm I$  lies on a unique longitude,

The elements  $\pm I$  lie on every longitude.

Therefore, every great circle in  $SU_2$  is a coset of  $L = W \cap S^3$ .

Hence, **every great circle in  $SU_2$  is a coset of one of the longitudes.**

4. a

Consider the centralizer of  $j$  in  $SU_2$ .

Where,

$$j = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Assume,

$$P \in SU_2$$

Then,

$$P = \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix}$$

Where,

$$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1.$$

Now,

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix} = \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

After matrix multiplication,

$$\begin{bmatrix} -x_2 + x_3 i & x_0 - x_1 i \\ -x_0 - x_1 i & -x_2 - x_3 i \end{bmatrix} = \begin{bmatrix} -x_2 - x_3 i & x_0 + x_1 i \\ -x_0 + x_1 i & -x_2 + x_3 i \end{bmatrix}$$

Therefore,

On comparing both sides,

$$x_3 = 0$$

$$x_1 = 0$$

So,

After putting these values in  $x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$ ,

As a result,

$$x_0^2 + x_2^2 = 1.$$

Thus,

The centralizer of  $j$  in  $SU_2$  is  $\begin{bmatrix} x_0 & x_2 \\ -x_2 & x_0 \end{bmatrix}$  where  $x_0^2 + x_2^2 = 1$ .

Hence, **the centralizer of  $j$  in  $SU_2$  is  $SO_2$ .**

## Section 4

1. a

Consider  $W$  be the space of real skew-symmetric  $3 \times 3$  matrices and the operation  $P * A = PAP^T$  of  $SO_3$  on  $W$ .

Since,

$$P \in SO_3$$

And,

$$A \in W$$

Then,

The operation,

$$P * A = PAP^T$$

$$\begin{aligned}
 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & \sin \theta & -\cos \theta \end{bmatrix} \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & \sin \theta & -\cos \theta \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & \sin \theta & -\cos \theta \end{bmatrix} \begin{bmatrix} 0 & -a \cos \theta - b \sin \theta & -a \sin \theta + b \cos \theta \\ a & -c \sin \theta & c \cos \theta \\ b & c \cos \theta & c \sin \theta \end{bmatrix} \\
 &= \begin{bmatrix} 0 & -a \cos \theta - b \sin \theta & -a \sin \theta + b \cos \theta \\ a \cos \theta + b \sin \theta & 0 & c \\ a \sin \theta - b \cos \theta & -c & 0 \end{bmatrix}
 \end{aligned}$$

Therefore, the matrix obtained by the operation  $P * A$  is also real skew-symmetric  $3 \times 3$  matrices.

Hence, **the orbit for the operation  $P * A = PAP^T$  of  $SO_3$  on  $W$  is also  $W$ .**

2. a

Consider the rotation group  $SO_3$  mapped to a 2-sphere by sending a rotation matrix to its first column.

[Comment](#)

Step 2 of 2 ^

Since,

The fibre of this map will consist of the special unitary group  $SU_2$ .

As,

Every element of  $SU_2$  can be described as a rotation together with a choice of spin.

Hence, **the fibre of the map is spin group  $SU_2$ .**

3. a



Consider the orthogonal representation  $\varphi: SU_2 \rightarrow SO_3$ .

[Comment](#)

Step 2 of 2 ^

Since,

The homomorphism  $\phi: U_2 \rightarrow SO_3$  can be considered as spin homomorphism with contraction from  $n$ -manifold to  $1$ -manifold.

And,

The kernel of  $\phi$  will be spin group  $SU_2$ .

Therefore, the homomorphism  $\phi: U_2 \rightarrow SO_3$  is spin homomorphism with  $n$ -manifold to  $1$ -manifold contraction and kernel of  $\phi$  is spin group  $SU_2$ .

4. a

Consider the matrix of rotation  $\gamma_p$ .

[Comment](#)

Step 2 of 3 ^

(a)

Since,

The matrix of rotation  $\gamma_p$

$$= (\cos \theta)I + (\sin \theta)A$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

And,

$$\text{trace} = 1 + \cos \theta + \cos \theta$$

$$= 1 + 2 \cos \theta$$

Therefore, the matrix of rotation  $\gamma_p$  is  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$  and its trace is  $1 + 2 \cos \theta$ .

(b)

Consider,

$$X = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},$$

$$Y = \begin{bmatrix} 0 & \cos \theta & \sin \theta \end{bmatrix},$$

$$Z = \begin{bmatrix} 0 & -\sin \theta & \cos \theta \end{bmatrix}.$$

As,

$$XY = YZ = ZX = 0$$

Therefore, the matrix of rotation  $\gamma_p$  is orthogonal.

5. a

Consider the rotation through conjugation by an element of  $SU_2$ .

[Comment](#)

Step 2 of 2 ^

Since,

The equator  $E$  is the unit 2-sphere in the three-dimensional space  $V$  of trace zero, skew Hermitian matrices.

And,

The latitudes in  $SU_2$  are conjugacy classes. For a given  $c$  in the interval  $-1 < c < 1$ , the latitude  $\{x_0 = c\}$  consists of the matrices  $P$  in  $SU_2$  such that  $\text{trace} P = 2c$ .

As,

The conjugation by an element  $P$  of  $SU_2$  preserves both the trace and the skew-Hermitian property.

So,

The conjugation  $\gamma_P$  operates on the whole space  $V$ .

Also,

The operator  $\gamma_P$  is a rotation of  $E$  and  $V$ .

Therefore, **the conjugation by an element of the unitary group  $SU_2$  rotates latitudes.**

6. a

**Conjugacy class:**

Suppose an element  $g$  of a group  $G$  then its conjugacy class defined by,

$$C_g = \{xgx^{-1} : x \in G\}$$

To describe the conjugacy classes in  $SO_3$ ;

[Comment](#)

Step 2 of 4 ^

(a)

Consider the following matrices:

$$A = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

And,

$$C = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}$$

Where,  $A, B, C$  belongs to  $SO_3$  and elements of  $SO_3$  operates on  $\mathbb{R}^3$  as rotations.

**Hence,** matrix  $A$  forms a conjugacy class of  $\mathbb{R}^3$  because these are rotations in the  $xy$ -plane and matrices of this form are commutative.

(b)

The spin homomorphism can be defined by a surjective homomorphism  $\gamma: SU_2 \rightarrow SO_3$  and its kernel is the center  $\{\pm I\}$  of  $SU_2$ .

Consider the  $2 \times 2$  matrix  $P$  in  $SU_2$ , say

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

Define a map  $\gamma: SU_2 \rightarrow SO_3$  by,

$$\gamma\left(\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}\right) = H$$

Where, for all  $3 \times 3$  matrix  $H$  in  $SO_3$  and every element or matrix of  $SO_3$  obtain by the rotation of axis.

So, by the definition of spin homomorphism every element of  $SU_2$  except  $\pm I$  can be described as a nontrivial rotation with a different choice of spin and  $SU_2$  is called spin group and

**Hence**, the spin homomorphism  $SU_2 \rightarrow SO_3$  is related to the conjugacy class in the two groups.

(c)

To describe the conjugacy classes in  $SO_3$ ;

Since the conjugacy classes in  $SU_2$  are spheres and every element of  $SU_2$  is conjugate to an element of the form:

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

Consider the matrices;

$$A = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

And,

$$C = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}$$

Matrix  $A, B, C$  as given in part (a) formed by a rotation of axis  $xy, yz$  and  $zx$ -plane of  $SO_3$ .

Since elements of  $SO_3$  corresponds to pair of antipodal points of  $SU_2$ , so obtain a group  $SO_3$  topologically by identify the antipodal points on the 3-sphere and  $SO_3$  is homeomorphic to projective 3-space  $\mathbb{R}P^3$ .

**Hence**, the conjugacy class of  $SO_3$  obtained by the rotation of any axis and conjugacy classes in  $SO_3$  are in 3-sphere.

7. a

Let  $A$  be a  $m \times n$  matrix. Left multiplication by a matrix  $A$  is defined as follows-

Let  $P$  be a matrix of order  $n \times p$  the left multiplication of  $P$  by  $A$  is given by the matrix  $AP$ .

(a)

Let  $P$  be a fixed matrix in  $SU_2$ .

Then  $P$  has the form

$$P = \begin{bmatrix} a & b \\ -\bar{b} & a \end{bmatrix}, \text{ where } a, b \in \mathbb{C} \text{ such that } a = x_0 + ix_1, \quad b = x_2 + ix_3 \text{ and } a\bar{a} + b\bar{b} = 1$$

Let  $R$  be any matrix in  $SU_2$ .

Then,

$$R = \begin{bmatrix} p & q \\ -q & p \end{bmatrix}, \text{ where } p, q \in \mathbb{C} \text{ such that } p = a_0 + ia_1, \quad q = a_2 + ia_3 \text{ and } p\bar{p} + q\bar{q} = 1$$

Consider the matrix product  $PR$

$$\begin{aligned} PR &= \begin{bmatrix} a & b \\ -\bar{b} & a \end{bmatrix} \begin{bmatrix} p & q \\ -q & p \end{bmatrix} \\ &= \begin{bmatrix} ap - b\bar{q} & aq + b\bar{p} \\ -(aq + b\bar{p}) & (ap - b\bar{q}) \end{bmatrix} \end{aligned}$$

Let  $c = ap - b\bar{q}$  and  $d = aq + b\bar{p}$ , so

$$\begin{aligned} c &= ap - b\bar{q} \\ &= (x_0 + ix_1)(a_0 + ia_1) - (x_2 + ix_3)(a_2 - ia_3) \\ &= a_0x_0 - a_1x_1 + a_3x_3 - a_2x_2 + i(a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3) \end{aligned}$$

$$\begin{aligned} d &= aq + b\bar{p} \\ &= (x_0 + ix_1)(a_2 + ia_3) + (x_2 + ix_3)(a_0 - ia_1) \\ &= a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3 + i(a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3) \end{aligned}$$

Now let  $c = z_0 + iz_1$  and  $d = z_2 + iz_3$

Now compare the terms on both sides

$$\begin{aligned} z_0 &= a_0x_0 - a_1x_1 + a_3x_3 - a_2x_2 \\ z_1 &= a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3 \\ z_2 &= a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3 \\ z_3 &= a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3 \end{aligned}$$

Re write the above system of equation in matrix form.

Let the coefficient matrix be denoted by  $Q$ .

Thus in matrix form the system of equation becomes

$$\begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} a_0 & -a_1 & -a_2 & a_3 \\ a_1 & a_0 & a_3 & -a_2 \\ a_2 & -a_3 & a_0 & a_1 \\ a_3 & a_2 & -a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Here,

$$Q = \begin{bmatrix} a_0 & -a_1 & -a_2 & a_3 \\ a_1 & a_0 & a_3 & -a_2 \\ a_2 & -a_3 & a_0 & a_1 \\ a_3 & a_2 & -a_1 & a_0 \end{bmatrix}$$

Now transpose of the matrix  $Q$  is given by

$$Q' = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ -a_1 & a_0 & -a_3 & a_2 \\ -a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{bmatrix}$$

Evaluate  $QQ'$  and  $Q'Q$

$$\begin{aligned} QQ' &= \begin{bmatrix} a_0 & -a_1 & -a_2 & a_3 \\ a_1 & a_0 & a_3 & -a_2 \\ a_2 & -a_3 & a_0 & a_1 \\ a_3 & a_2 & -a_1 & a_0 \end{bmatrix} \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ -a_1 & a_0 & -a_3 & a_2 \\ -a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Similarly  $Q'Q = I_4$

So the matrix  $Q$  is orthogonal.

**Therefore, the result in the question has been proved.**

(b)

Let  $v = (x_0, x_1, x_2, x_3)$  and  $w = (x_0', x_1', x_2', x_3')$ .

Consider the vector  $Qv$

$$Qv = \begin{bmatrix} a_0 & -a_1 & -a_2 & a_3 \\ a_1 & a_0 & a_3 & -a_2 \\ a_2 & -a_3 & a_0 & a_1 \\ a_3 & a_2 & -a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Now multiply the above matrixes in order to obtain a matrix of  $4 \times 1$  order

$$Qv = \begin{bmatrix} a_0x_0 - a_1x_1 - a_2x_2 + a_3x_3 \\ a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3 \\ a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3 \\ a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3 \end{bmatrix}$$

Similarly evaluate the vector  $Qw$  which is given by

$$\begin{aligned} Qw &= \begin{bmatrix} a_0 & -a_1 & -a_2 & a_3 \\ a_1 & a_0 & a_3 & -a_2 \\ a_2 & -a_3 & a_0 & a_1 \\ a_3 & a_2 & -a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_0' \\ x_1' \\ x_2' \\ x_3' \end{bmatrix} \\ &= \begin{bmatrix} a_0x_0' - a_1x_1' - a_2x_2' + a_3x_3' \\ a_1x_0' + a_0x_1' + a_3x_2' - a_2x_3' \\ a_2x_0' - a_3x_1' + a_0x_2' + a_1x_3' \\ a_3x_0' + a_2x_1' - a_1x_2' + a_0x_3' \end{bmatrix} \end{aligned}$$

Now evaluate the dot product of the vectors  $Qv$  and  $Qw$

$$\begin{aligned} Qv \cdot Qw &= \begin{bmatrix} a_0x_0 - a_1x_1 - a_2x_2 + a_3x_3 \\ a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3 \\ a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3 \\ a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3 \end{bmatrix} \cdot \begin{bmatrix} a_0x_0' - a_1x_1' - a_2x_2' + a_3x_3' \\ a_1x_0' + a_0x_1' + a_3x_2' - a_2x_3' \\ a_2x_0' - a_3x_1' + a_0x_2' + a_1x_3' \\ a_3x_0' + a_2x_1' - a_1x_2' + a_0x_3' \end{bmatrix} \\ &= (a_0x_0 - a_1x_1 + a_3x_3 - a_2x_2)(a_0x_0' - a_1x_1' - a_2x_2' + a_3x_3') + \\ &\quad (a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3)(a_1x_0' + a_0x_1' + a_3x_2' - a_2x_3') + \\ &\quad (a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3)(a_2x_0' - a_3x_1' + a_0x_2' + a_1x_3') + \\ &\quad (a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3)(a_3x_0' + a_2x_1' - a_1x_2' + a_0x_3') \end{aligned}$$

Now evaluate the dot product of the vectors  $\underline{Q}_w$  and  $\underline{Q}_v$

$$\begin{aligned}\underline{Q}_v \cdot \underline{Q}_w &= \begin{bmatrix} a_0x_0 - a_1x_1 + a_3x_3 - a_2x_2 \\ a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3 \\ a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3 \\ a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3 \end{bmatrix} \cdot \begin{bmatrix} a_0x_0' - a_1x_1' - a_2x_2' + a_3x_3' \\ a_1x_0' + a_0x_1' + a_3x_2' - a_2x_3' \\ a_2x_0' - a_3x_1' + a_0x_2' + a_1x_3' \\ a_3x_0' + a_2x_1' - a_1x_2' + a_0x_3' \end{bmatrix} \\ &= (a_0x_0 - a_1x_1 + a_3x_3 - a_2x_2)(a_0x_0' - a_1x_1' - a_2x_2' + a_3x_3') + \\ &\quad (a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3)(a_1x_0' + a_0x_1' + a_3x_2' - a_2x_3') + \\ &\quad (a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3)(a_2x_0' - a_3x_1' + a_0x_2' + a_1x_3') + \\ &\quad (a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3)(a_3x_0' + a_2x_1' - a_1x_2' + a_0x_3')\end{aligned}$$

Now on simplify the term on the right hand side of the above equation

$$\begin{aligned}\underline{Q}_v \cdot \underline{Q}_w &= (a_0x_0 - a_1x_1 + a_3x_3 - a_2x_2)(a_0x_0' - a_1x_1' - a_2x_2' + a_3x_3') + \\ &\quad (a_1x_0 + a_0x_1 + a_3x_2 - a_2x_3)(a_1x_0' + a_0x_1' + a_3x_2' - a_2x_3') + \\ &\quad (a_2x_0 - a_3x_1 + a_0x_2 + a_1x_3)(a_2x_0' - a_3x_1' + a_0x_2' + a_1x_3') + \\ &\quad (a_3x_0 + a_2x_1 - a_1x_2 + a_0x_3)(a_3x_0' + a_2x_1' - a_1x_2' + a_0x_3') \\ &= x_0x_0' + x_1x_1' + x_2x_2' + x_3x_3'\end{aligned}$$

Now the term on the right hand side is the dot product of the vectors  $v$  and  $w$

Consider the dot product of the vectors  $v$  and  $w$

$$\begin{aligned}v \cdot w &= (x_0, x_1, x_2, x_3) \cdot (x_0', x_1', x_2', x_3') \\ &= x_0x_0' + x_1x_1' + x_2x_2' + x_3x_3'\end{aligned}$$

Thus  $v \cdot w = \underline{Q}_v \cdot \underline{Q}_w$  for every  $v, w$

**Therefore, the result in the question has been proved.**

8. a

Let  $\mathcal{W}$  be the real vector space of Hermitian  $2 \times 2$  matrices.

[Comment](#)

Step 2 of 8 ^

(a)

To show that the rule  $P \cdot A = PAP^*$  defines an operation of  $SL_2(\mathbb{C})$  on  $\mathcal{W}$ :

Dimension of  $SL_2(\mathbb{C})$  is 6 matrix and element in  $SL_2(\mathbb{C})$  can be defined by,

$$SL_2(\mathbb{C}) = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{C} \text{ and } \det A = 1 \right\}$$

Now multiply by  $P$  on the above,

$$SL_2(\mathbb{C}) = \left\{ PA = P \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{C} \text{ and } \det A = 1 \right\}$$

And a real Hermitian matrix has real eigenvalue so suppose  $P = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$  where  $x, y$  in  $\mathbb{R}$ .

Then, this defines an operation of  $SL_2(\mathbb{C})$  on  $\mathcal{W}$ :

**Hence,** the rule  $P \cdot A = PAP^*$  defines an operation of  $SL_2(\mathbb{C})$  on  $\mathcal{W}$ :



(b)

To show that the function  $\langle A, A' \rangle = \det(A + A') - \det A - \det A'$  is a bilinear form on  $W$ ;

Let  $A, B$  in  $W$  then,

$$\langle A + B, A' \rangle = \det((A + B) + A') - \det(A + B) - \det A'$$

Apply  $\det(A + B) = \det A + \det B + \text{tr}(A'B)$  in the above,

$$\begin{aligned}\langle A + B, A' \rangle &= \det(A + B) + \det A' + \text{tr}((A + B)'A') - \det(A + B) - \det A' \\ &= \text{tr}((A' + B')A') \\ &= \text{tr}(A'A' + B'A')\end{aligned}$$

Apply  $\text{tr}(A + B) = \text{tr}A + \text{tr}B$  in the above,

$$\begin{aligned}\langle A + B, A' \rangle &= \text{tr}(A'A') + \text{tr}(B'A') \\ &= \det(A + A') - \det A - \det A' + \det(B + A') - \det B - \det A' \\ &= \langle A, A' \rangle + \langle B, A' \rangle\end{aligned}$$

Now, to show that  $\langle A, A' + B' \rangle = \langle A, A' \rangle + \langle A, B' \rangle$ ;

Let  $B'$  in  $W$  then,

$$\langle A, A' + B' \rangle = \det(A + (A' + B')) - \det A - \det(A' + B')$$

Apply  $\det(A + B) = \det A + \det B + \text{tr}(A'B)$  in the above,

$$\begin{aligned}\langle A, A' + B' \rangle &= \det(A) + \det(A' + B') + \text{tr}(A'(A' + B')) - \det A - \det(A' + B') \\ &= \text{tr}(A'(A' + B')) \\ &= \text{tr}(A'A' + A'B')\end{aligned}$$

Apply  $\text{tr}(A + B) = \text{tr}A + \text{tr}B$  in the above,

$$\begin{aligned}\langle A, A' + B' \rangle &= \text{tr}(A'A') + \text{tr}(A'B') \\ &= \det(A + A') - \det A - \det A' + \det(A + B') - \det A - \det B' \\ &= \langle A, A' \rangle + \langle A, B' \rangle\end{aligned}$$

Now, to show that  $\langle \lambda A, A' \rangle = \lambda \langle A, A' \rangle$ ;

$$\langle \lambda A, A' \rangle = \det((\lambda A) + A') - \det(\lambda A) - \det A'$$

Apply  $\det(A + B) = \det A + \det B + \text{tr}(A'B)$  in the above,

$$\begin{aligned}\langle \lambda A, A' \rangle &= \det(\lambda A) + \det(A') + \text{tr}((\lambda A)'A') - \det(\lambda A) - \det A' \\ &= \text{tr}((\lambda A)'A') \\ &= \text{tr}(\lambda A'A')\end{aligned}$$

Apply  $\text{tr}(\lambda A) = \lambda \text{tr}A$  in the above,

$$\begin{aligned}\langle \lambda A, A' \rangle &= \lambda \text{tr}(A'A') \\ &= \lambda \langle A, A' \rangle\end{aligned}$$

Hence, the function  $\langle A, A' \rangle = \det(A + A') - \det A - \det A'$  is a bilinear form on  $W$ ;

To show that signature is  $(3,1)$ .

It is known that signature of  $n \times n$  is  $\left(\frac{n^2+n}{2}, \frac{n^2-n}{2}\right)$ .

Here,  $W$  is the real vector space of Hermitian  $2 \times 2$  matrices. Substitute  $n = 2$  into the above, then signature is:

$$\begin{aligned} \left(\frac{n^2+n}{2}, \frac{n^2-n}{2}\right) &= \left(\frac{2^2+2}{2}, \frac{2^2-2}{2}\right) \\ &= \left(\frac{4+2}{2}, \frac{4-2}{2}\right) \\ &= \left(\frac{6}{2}, \frac{2}{2}\right) \\ &= (3,1) \end{aligned}$$

Hence, its signature is  $(3,1)$ .

[Comment](#)

Step 8 of 8 ^

(c)

To define a homomorphism  $\varphi: SL_2(\mathbb{C}) \rightarrow O_{3,1}$ , whose kernel is  $\{\pm I\}$ :

$$\varphi \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

Where, matrix  $A$  in  $O_{3,1}$  has a polar decomposition of the form:

$$A = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^T} & v \\ v^T & c \end{pmatrix}$$

Or,

$$A = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^T} & v \\ v^T & c \end{pmatrix}$$

Where  $Q \in O_3$  and  $c = \sqrt{\|v\|^2 + 1}$

And here  $\det(A) = \pm 1$

Hence, a homomorphism  $\varphi: SL_2(\mathbb{C}) \rightarrow O_{3,1}$ , whose kernel is  $\{\pm I\}$ :

$$\boxed{\varphi \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A}$$

9. a

(a)

To show that every element of  $SO_3$  can be written as a product of  $ABA'$  where  $A$  and  $A'$  are in  $H_1$  and  $B$  is in  $H_2$ .

Consider the subgroup of  $SO_3$  are  $H_i$  of rotations about the  $x_i$  axis, where  $i = 1, 2, 3$ .

Group  $SO_3$  is defined as  $\{A \in SO_3 : AA' = I \text{ and } \det A = 1\}$

Suppose that  $A$  and  $A'$  are in  $H_1$  and counter clockwise rotation about the positive  $x_3$ -axis by angle  $\theta$  as shown below:

$$A = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrix  $A'$  is:

$$A' = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Suppose that  $B$  is in  $H_2$  and counter clockwise rotation about the positive  $x_2$ -axis by angle  $\theta$  as shown below:

$$B = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}$$

Now,  $B$  and  $A$  matrix are commute, that is  $AB = BA$ , then

$$\begin{aligned} ABA' &= BAA' \\ &= BI \\ &= B \end{aligned}$$

Where,  $B$  is in  $H_2$  and  $H_2$  is subgroup of  $SO_3$ . So  $B$  is in  $SO_3$ .

**Hence**, every element of  $SO_3$  can be written as a product of  $ABA'$ .

(b)

**Double cosets:**

Let  $H$  and  $K$  be subgroups of a group  $G$  and let  $g$  be an element of  $G$  the set  $HgK = \{x \in G \mid x = h g k \text{ for some } h \in H, k \in K\}$  is called a double coset.

To describe the double cosets  $H_1QH_1$ ;

By the definition of the double cosets, here  $H_1$  is subgroup of a group  $SO_3$  and let  $Q$  be a matrix of  $SO_3$  with determinant is one.

Thus, define a double coset  $H_1QH_1 = \{A \in SO_3 \mid A = BQC \text{ for some } B, C \in H_1\}$ .

## Section 5

1. a

Consider the image of a one-parameter group in  $GL_n$ .

[Comment](#)

---

Step 2 of 2 ^

Since,

The one-parameter group is bijective correspondence with  $n \times n$  matrices in  $GL_n$ .

Therefore,

The image of a one-parameter group in  $GL_n$  cannot cross itself due to one-to-one mapping.

Hence, **the image of a one-parameter group in  $GL_n$  does not cross itself.**

2. a

Consider the one-parameter groups in  $U_2$ .

[Comment](#)

---

Step 2 of 2 ^

Since,

If the matrix  $A$  is complex skew-Hermitian  $2 \times 2$  matrix, so is  $tA$  and  $e^{tA}$  is unitary for all  $t$ .

So,

The homomorphisms  $t \rightarrow e^{tA}$  are one-parameter groups in the unitary group  $U_2$ .

Hence, **the one-parameter groups in the unitary group  $U_2$  are the homomorphisms  $t \rightarrow e^{tA}$ , where  $A$  is a complex skew-Hermitian  $2 \times 2$  matrix.**

3. a

A one parameter group is a differentiable homomorphism  $\varphi: \mathbb{R}^+ \rightarrow G$  defined by

$\varphi(t) = e^{tA}$ , where  $A$  is a matrix

In layman terms one must evaluate all the matrices such that  $e^{tA}$  belongs to  $G$  for all  $t \in \mathbb{R}$

Let  $G$  denote the group of all  $2 \times 2$  real, invertible and diagonal matrix.

In mathematical form  $G$  is the group of matrix  $A$  of the form

$$A = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, \text{ where } x, y \in \mathbb{R} - \{0\}$$

Since  $x, y \in \mathbb{R} - \{0\}$  thus the matrix is invertible.

Let  $A$  be a matrix such that  $e^{tA} \in G \quad \forall t \in \mathbb{R}$ .

Let  $r(t)_A = e^{tA}$

Now let  $r(t)_A \in G$  for every  $t$ .

Thus  $r(t)_A$  has the form

$$r(t)_A = \begin{pmatrix} x(t) & 0 \\ 0 & y(t) \end{pmatrix}$$

Now since  $e^{tA}$  is differentiable so  $p(t)_A$  is also differentiable

$$r(t)_A = \begin{pmatrix} x(t) & 0 \\ 0 & y(t) \end{pmatrix}$$

On differentiating both sides with respect to  $t$

$$\frac{d}{dt}(r(t)_A) = \begin{pmatrix} \frac{d(x(t))}{dt} & 0 \\ 0 & \frac{d(y(t))}{dt} \end{pmatrix}$$

Since  $r(t)_A = e^{tA}$ , differentiate both sides to obtain

$$\begin{aligned} \frac{d}{dt}(r(t)_A) &= \frac{d(e^{tA})}{dt} \\ &= Ae^{tA} \end{aligned}$$

Now  $\frac{d}{dt}(r(t)_A)$  at  $t=0$  is given by

$$\begin{aligned} \left. \frac{d}{dt}(r(t)_A) \right|_{t=0} &= \begin{pmatrix} \frac{d(x(t))}{dt} & 0 \\ 0 & \frac{d(y(t))}{dt} \end{pmatrix}_{t=0} \\ &= \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \end{aligned}$$

At  $t=0$ ,  $\left. \frac{d}{dt}(r(t)_A) \right|_{t=0}$  is also given by

$$\begin{aligned} \left. \frac{d}{dt}(r(t)_A) \right|_{t=0} &= (Ae^{tA})_{t=0} \\ &= A \end{aligned}$$

So the matrix  $A$  is of the form

$$A = \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}, \text{ where } p = \left( \frac{dx}{dt} \right)_{t=0} \text{ and } q = \left( \frac{dy}{dt} \right)_{t=0}$$

Let  $A$  be a matrix of the above derived form

Now evaluate  $e^{tA}$  for such matrices

$$e^{tA} = I + tA + \frac{t^2}{2!}A^2 + \dots + \frac{t^n}{n!}A^n + \dots$$

For the matrix  $A$ , its successive powers are given by

$$\begin{aligned} A^2 &= \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \\ &= \begin{pmatrix} p^2 & 0 \\ 0 & q^2 \end{pmatrix} \end{aligned}$$

Similarly,

$$\begin{aligned} A^3 &= A^2 A \\ &= \begin{pmatrix} p^2 & 0 \\ 0 & q^2 \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \\ &= \begin{pmatrix} p^3 & 0 \\ 0 & q^3 \end{pmatrix} \end{aligned}$$

In general terms

$$\begin{aligned} A^n &= A^{n-1} A \\ &= \begin{pmatrix} p^{n-1} & 0 \\ 0 & q^{n-1} \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \\ &= \begin{pmatrix} p^n & 0 \\ 0 & q^n \end{pmatrix} \end{aligned}$$

Now, substitute the value of the matrix powers in the expansion of  $e^{tA}$

$$\begin{aligned} e^{tA} &= I + tA + \frac{t^2}{2!}A^2 + \dots + \frac{t^n}{n!}A^n + \dots \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} + \frac{t^2}{2!} \begin{pmatrix} p^2 & 0 \\ 0 & q^2 \end{pmatrix} + \dots + \frac{t^n}{n!} \begin{pmatrix} p^n & 0 \\ 0 & q^n \end{pmatrix} + \dots \end{aligned}$$

Add all the terms by first multiplying the matrix with the corresponding scalars

The following matrix is obtained after the series of above procedure

$$e^{tA} = \begin{pmatrix} 1 + tp + \dots + \frac{t^n p^n}{n!} + \dots & 0 \\ 0 & 1 + tq + \dots + \frac{t^n q^n}{n!} + \dots \end{pmatrix}$$

Also since  $e^{tA} \in G$  so

$$e^{tA} = \begin{pmatrix} x(t) & 0 \\ 0 & y(t) \end{pmatrix},$$

Now compare the terms of both the matrix

$$\begin{pmatrix} x(t) & 0 \\ 0 & y(t) \end{pmatrix} = \begin{pmatrix} 1 + tp + \dots + \frac{t^n p^n}{n!} + \dots & 0 \\ 0 & 1 + tq + \dots + \frac{t^n q^n}{n!} + \dots \end{pmatrix}$$

So,

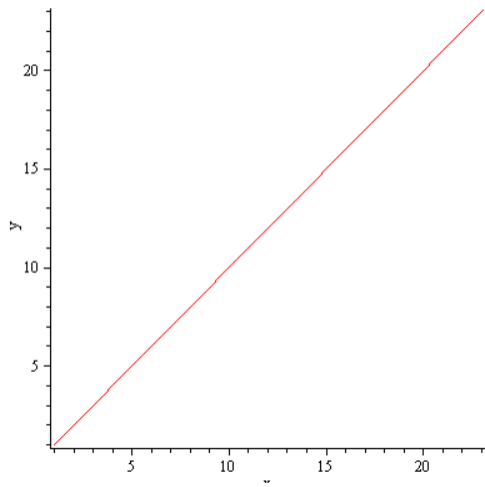
$$x(t) = 1 + tp + \dots + \frac{t^n p^n}{n!} + \dots, \text{ and}$$

$$y(t) = 1 + tq + \dots + \frac{t^n q^n}{n!} + \dots$$

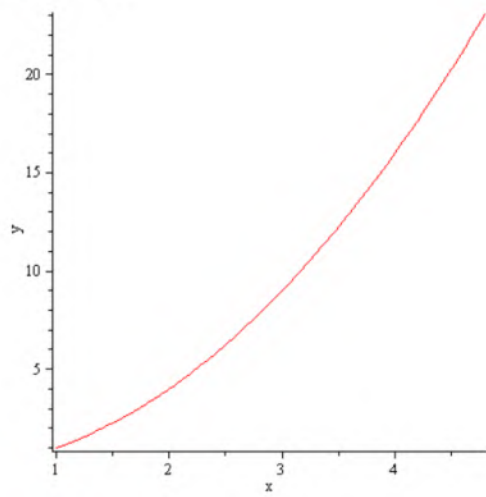
Clearly by the expansion formula  $x(t) = e^{pt}$  and  $y(t) = e^{qt}$ .

For  $p, q = 1$

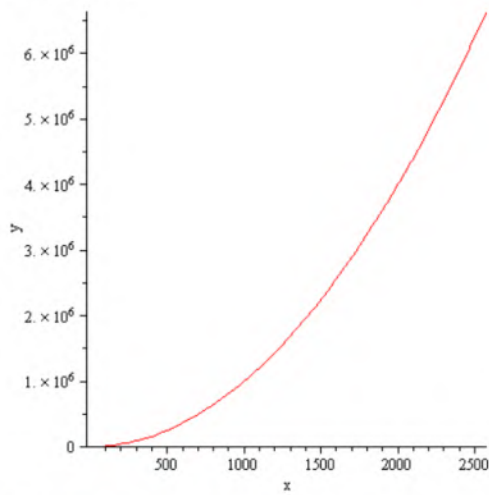




For  $p = 0.5, q = 1$



For  $p = 2.5, q = 5$



Therefore, the result mentioned in the question has been proved.

4. a

Consider the conditions on a matrix  $A$  so that  $e^{tA}$  is a one-parameter group.

[Comment](#)

Step 2 of 3 ^

(a)

Consider, the special unitary group  $SU_n$ ,

Since,

$$e^{\text{trace} A} = \det e^A$$

Then,

$$\text{trace}(A) = 0$$

And,

The matrix  $A$  is a complex skew-Hermitian matrix.

Therefore, **the matrix  $A$  has trace zero and complex skew-Hermitian matrix.**

(b)

Consider the Lorentz group  $O_{3,1}$ ,

Since,

$$O_{3,1} = \{P \in GL_4 : P^T I_{3,1} P = I_{3,1}\}$$

Then,

$$A + A^T = 0$$

Therefore, **matrix  $A$  is skew symmetric matrix.**

5. a

A one parameter group is a differentiable homomorphism  $\varphi: \mathbb{R}^+ \rightarrow G$  defined by

$$\varphi(t) = e^{tA}, \text{ where } A \text{ is a matrix}$$

In layman terms one must evaluate all the matrices such that  $e^{tA}$  belongs to  $G$  for all  $t \in \mathbb{R}$

[Comment](#)

Step 2 of 5 ^

(a)

Let  $G$  be the group of real matrices of the form

$$G = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}, \text{ where } x > 0$$

Let  $A$  be a matrix such that  $e^{tA} \in G \quad \forall t \in \mathbb{R}$ .

$$\text{Let } r(t)_A = e^{tA}$$

Now let  $r(t)_A \in G$  for every  $t$ .

Thus  $r(t)_A$  has the form

$$r(t)_A = \begin{pmatrix} x(t) & y(t) \\ 0 & 1 \end{pmatrix}$$

Now since  $e^{tA}$  is differentiable so  $p(t)_A$  is also differentiable

$$r(t)_A = \begin{pmatrix} x(t) & y(t) \\ 0 & 1 \end{pmatrix}$$

On differentiating both sides with respect to  $t$

$$\frac{d}{dt}(r(t))_A = \begin{pmatrix} \frac{d(x(t))}{dt} & \frac{d(y(t))}{dt} \\ 0 & 1 \end{pmatrix}$$

Since  $r(t)_A = e^{tA}$ , differentiate both sides to obtain

$$\begin{aligned} \frac{d}{dt}(r(t))_A &= \frac{d(e^{tA})}{dt} \\ &= Ae^{tA} \end{aligned}$$

Now  $\frac{d}{dt}(r(t))_A$  at  $t=0$  is given by

$$\begin{aligned} \left. \frac{d}{dt}(r(t))_A \right|_{t=0} &= \begin{pmatrix} \frac{d(x(t))}{dt} & \frac{d(y(t))}{dt} \\ 0 & 1 \end{pmatrix}_{t=0} \\ &= \begin{pmatrix} p & q \\ 0 & 0 \end{pmatrix} \end{aligned}$$

At  $t=0$ ,  $\left. \frac{d}{dt}(r(t))_A \right|_{t=0}$  is also given by

$$\begin{aligned} \left. \frac{d}{dt}(r(t))_A \right|_{t=0} &= (Ae^{tA})_{t=0} \\ &= A \end{aligned}$$

So the matrix  $A$  is of the form

$$A = \begin{pmatrix} p & q \\ 0 & 0 \end{pmatrix}, \text{ where } p = \left( \frac{dx}{dt} \right)_{t=0} \text{ and } q = \left( \frac{dy}{dt} \right)_{t=0}$$

**(b)**

Let  $A$  be a matrix of the form

$$A = \begin{pmatrix} p & q \\ 0 & 0 \end{pmatrix}, \text{ where } p = \left( \frac{dx}{dt} \right)_{t=0} \text{ and } q = \left( \frac{dy}{dt} \right)_{t=0}$$

Now evaluate  $e^{tA}$  for such matrices

$$e^{tA} = I + tA + \frac{t^2}{2!}A^2 + \dots + \frac{t^n}{n!}A^n + \dots$$

For the matrix  $A$ , its successive powers are given by

$$\begin{aligned} A^2 &= \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} p^2 & pq \\ 0 & 1 \end{pmatrix} \\ &= p \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Similarly,

$$\begin{aligned} A^3 &= A^2 A \\ &= p \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \\ &= p^2 \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \end{aligned}$$

In general terms

$$\begin{aligned} A^n &= A^{n-1} A \\ &= p^{n-2} \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \\ &= p^{n-1} \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Now, substitute the value of the matrix powers in the expansion of  $e^{tA}$

$$\begin{aligned}
e^{tA} &= I + tA + \frac{t^2}{2!} A^2 + \dots + \frac{t^n}{n!} A^n + \dots \\
&= I + tA + \frac{t^2 p}{2!} A + \dots + \frac{t^n p^{n-1}}{n!} A + \dots
\end{aligned}$$

Multiply and divide the expression  $tA + \frac{t^2 p}{2!} A + \dots + \frac{t^n p^{n-1}}{n!} A + \dots$  by  $p$

Then the expression  $\frac{1}{p} \left( tA + \frac{t^2 p^2}{2!} A + \dots + \frac{t^n p^n}{n!} A + \dots \right)$  can be simplified in the following way

$$\begin{aligned}
\frac{1}{p} \left( tA + \frac{t^2 p^2}{2!} A + \dots + \frac{t^n p^n}{n!} A + \dots \right) &= \frac{A}{p} \left( t + \frac{t^2 p^2}{2!} + \dots + \frac{t^n p^n}{n!} + \dots \right) \\
&= \frac{A}{p} (e^{pt} - 1)
\end{aligned}$$

Therefore, the matrix  $e^{tA}$  is given by  $\frac{A}{p} (e^{pt} - 1)$ , where  $p = \left( \frac{dx}{dt} \right)_{t=0}$ .

(c)

$$\text{Let } p = \left( \frac{dx}{dt} \right)_{t=0} \text{ and } q = \left( \frac{dy}{dt} \right)_{t=0}$$

Since the matrix  $e^{tA}$  is given by

$$e^{tA} = \frac{A}{p} (e^{pt} - 1), \text{ where } p = \left( \frac{dx}{dt} \right)_{t=0}$$

Also,

$$e^{tA} = \begin{pmatrix} x(t) & y(t) \\ 0 & 1 \end{pmatrix}$$

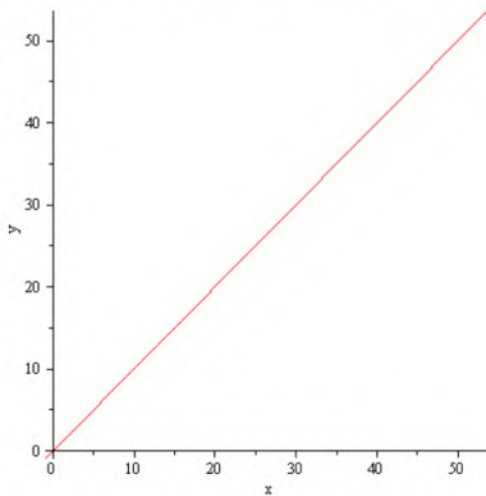
So the matrix must be equal, which implies

$$\begin{aligned}
\begin{pmatrix} x(t) & y(t) \\ 0 & 1 \end{pmatrix} &= \frac{A}{p} (e^{pt} - 1) \\
&= \frac{(e^{pt} - 1)}{p} \begin{pmatrix} p & q \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} e^{pt} - 1 & \frac{q}{p} (e^{pt} - 1) \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

This implies that  $x(t) = e^{pt} - 1$  and  $y(t) = \frac{q}{p} (e^{pt} - 1)$  for every  $t \in \mathbb{R}$

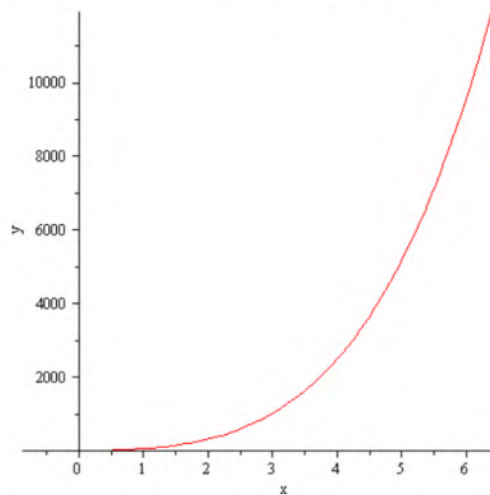
Now for  $p=1, q=1$ ,

Then,  $x(t) = e^t - 1$  and  $y(t) = e^t - 1$ , where  $t \in \mathbb{R}$



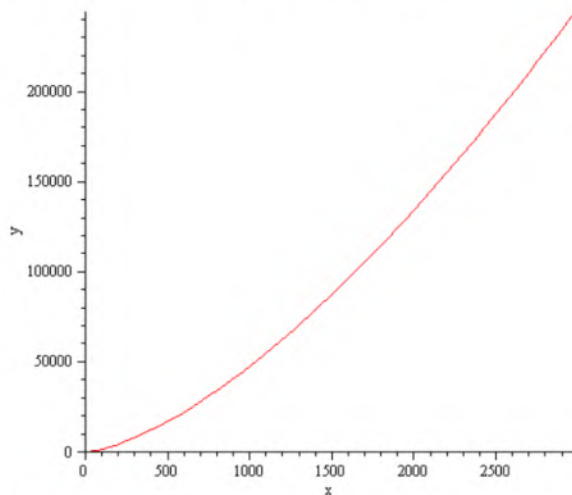
Now for  $p = 0.5, q = 2$ ,

Then,  $x(t) = e^{0.5t} - 1$  and  $y(t) = 4(e^{2t} - 1)$ , where  $t \in \mathbb{R}$



Now for  $p = 2, q = 3$

Then,  $x(t) = e^{2t} - 1$  and  $y(t) = \frac{3}{2}(e^{2t} - 1)$ , where  $t \in \mathbb{R}$



Therefore, the sketches of some of the one-parameter groups are as shown as above.

6. a

A one parameter group is a differentiable homomorphism  $\varphi: \mathbb{R}^+ \rightarrow G$  defined by

$$\varphi(t) = e^{tA}, \text{ where } A \text{ is a matrix}$$

In layman terms one must evaluate all the matrices such that  $e^{tA}$  belongs to  $G$  for all  $t \in \mathbb{R}$

To determine the conjugacy classes in  $G$  where  $G$  is the subgroup of matrices  $\begin{pmatrix} x & y \\ 0 & x^{-1} \end{pmatrix}$

Where,  $x > 0$  and  $y$  arbitrary

Suppose that matrix  $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$  is in subgroup  $G$ .

The conjugacy classes in  $G$  can be defined as,

$$CL(A) = \{BXB^{-1} : B \in G\}$$

$$\begin{aligned} BXB^{-1} &= \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ 0 & x^{-1} \end{pmatrix} \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & -bx + y \\ 0 & x^{-1} \end{pmatrix} \\ &= \begin{pmatrix} x & -bx + y + bx^{-1} \\ 0 & x^{-1} \end{pmatrix} \\ &= \begin{pmatrix} x & c \\ 0 & x^{-1} \end{pmatrix} \end{aligned}$$

Where,  $x > 0$  and  $-bx + y + bx^{-1} = c$  in  $\mathbb{R}$

And determinant of the above matrix is invertible. So it is form a conjugacy class of  $G$

Thus, conjugacy classes in  $G$  is  $\left\{ \begin{pmatrix} x & c \\ 0 & x^{-1} \end{pmatrix} : x > 0, c \in \mathbb{R} \right\}$ .

Let  $G$  be the subgroup of  $GL_2$  of matrices:

$$G = \begin{pmatrix} x & y \\ 0 & x^{-1} \end{pmatrix}$$

Where,  $x > 0$  and  $y$  arbitrary

Let  $A$  be a matrix such that  $e^{tA} \in G \quad \forall t \in \mathbb{R}$ .

Let  $r(t)_A = e^{tA}$

Now, let  $r(t)_A \in G$  for every  $t$ .

Thus,  $r(t)_A$  has the form:

$$r(t)_A = \begin{pmatrix} x(t) & y(t) \\ 0 & x^{-1}(t) \end{pmatrix}$$

Now since  $e^{tA}$  is differentiable, so  $p(t)_A$  is also differentiable.

$$r(t)_A = \begin{pmatrix} x(t) & y(t) \\ 0 & x^{-1}(t) \end{pmatrix}$$

On differentiating both sides with respect to  $t$

$$\frac{d}{dt}(r(t)_A) = \begin{pmatrix} \frac{d(x(t))}{dt} & \frac{d(y(t))}{dt} \\ 0 & \frac{d(x^{-1}(t))}{dt} \end{pmatrix}$$

Since  $r(t)_A = e^{tA}$ , differentiate both sides to obtain

$$\begin{aligned} \frac{d}{dt}(r(t)_A) &= \frac{d(e^{tA})}{dt} \\ &= Ae^{tA} \end{aligned}$$



Now  $\frac{d}{dt}(r(t)_A)$  at  $t=0$  is given by

$$\left. \frac{d}{dt}(r(t)_A) \right|_{t=0} = \begin{pmatrix} \frac{d(x(t))}{dt} & \frac{d(y(t))}{dt} \\ 0 & \frac{d(x^{-1}(t))}{dt} \end{pmatrix}_{t=0} = \begin{pmatrix} p & q \\ 0 & r \end{pmatrix}$$

At  $t=0$ ,  $\left. \frac{d}{dt}(r(t)_A) \right|_{t=0}$  is also given by

$$\left. \frac{d}{dt}(r(t)_A) \right|_{t=0} = (Ae^{tA})_{t=0} = A$$

Thus, the matrix  $A$  such that  $e^{tA}$  is a one-parameter group in  $G$  is of the form:

$$A = \begin{pmatrix} p & q \\ 0 & r \end{pmatrix}, \text{ where } p = \left( \frac{dx}{dt} \right)_{t=0}, q = \left( \frac{dy}{dt} \right)_{t=0} \text{ and } r = \frac{d(x^{-1}(t))}{dt}$$

7. a

Consider the one parameter groups in the group of invertible  $n \times n$  upper triangular matrices.

[Comment](#)

Step 2 of 2 ^

The invertible  $n \times n$  upper triangular matrix:

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots & \cdots & \cdots & \cdots & a_{1n} \\ & a_{22} & \cdots & \cdots & \cdots & \cdots & a_{2n} \\ & & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & \ddots & \ddots & \ddots & \vdots \\ & & & & \ddots & \ddots & \vdots \\ & & & & & \ddots & \vdots \\ & & & & & & a_{nn} \end{bmatrix}$$

Then,

$$e^{At} = \begin{bmatrix} e^{a_{11}t} & \cdots & \cdots & \cdots & \cdots & \cdots & e^{a_{1n}t} \\ & e^{a_{22}t} & \cdots & \cdots & \cdots & \cdots & e^{a_{2n}t} \\ & & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & \ddots & \ddots & \ddots & \vdots \\ & & & & \ddots & \ddots & \vdots \\ & & & & & \ddots & \vdots \\ & & & & & & e^{a_{nn}t} \end{bmatrix}$$

Therefore, the one parameter groups in the group of invertible  $n \times n$  upper triangular matrices is also invertible  $n \times n$  upper triangular matrices.

8. a

Consider  $\varphi(t) = e^{tA}$  be a one-parameter group in a subgroup  $G$  of  $GL_n$ .

[Comment](#)

Step 2 of 2 ^

Assume,  $P = \{e^{tA}, t \in \mathbb{R}^+\}$

$$X = PC, C \in GL_n$$

As,

$$\begin{aligned} \frac{dX}{dt} &= \frac{d}{dt}(PC) \\ &= \left[ \frac{d}{dt}(P = \{e^{tA}, t \in \mathbb{R}^+\}) \right] C \\ &= (AP)C \\ &= A(PC) \\ &= AX \end{aligned}$$

Hence, the cosets of image for one-parameter group are matrix solutions of the differential equation  $\frac{dX}{dt} = AX$ .

9. a

Consider  $\varphi(t) = e^{tA}$  be a one-parameter group in  $GL_n$ .

[Comment](#)

Step 2 of 2 ^

Assume,  $P = \{e^{tA}, t \in \mathbb{R}^+\}$

Then,  $\varphi(t)$  can be either  $\varphi(t) = e^{tA}$ ,  $\varphi(t) = e^{cAt}$  or  $\varphi(t) = e^{A \sin t}$ .

If,

$$\varphi(t) = e^{tA}$$

Then,  $\varphi(t)$  will be one-to-one mapping, then  $\text{Ker}(\varphi)$  will be trivial.

If,

$$\varphi(t) = e^{cAt}$$

Then,  $\varphi(t)$  will be constant mapping, then  $\text{Ker}(\varphi)$  will be whole group.

If,

$$\varphi(t) = e^{A \sin t}$$

Then,  $\varphi(t)$  will be periodic mapping, then  $\text{Ker}(\varphi)$  will be infinite cyclic group.

Hence, the group  $\text{Ker}(\varphi)$  is either trivial, or an infinite cyclic group, or the whole group.

10. a

Consider the differentiable homomorphisms from the circle group  $SO_2$  to  $GL_n$ .

Assume,

The mapping:

$$\phi: A\bar{A} \rightarrow A^{\frac{1}{n}}$$

Where,

$$n = 1, 2, 3, \dots$$

Consider,

$$A = \cos\theta + i\sin\theta$$

$$\bar{A} = \cos\theta - i\sin\theta$$

Then,

$$\begin{aligned} A\bar{A} &= (\cos\theta + i\sin\theta)(\cos\theta - i\sin\theta) \\ &= \cos^2\theta + \sin^2\theta \\ &= 1 \end{aligned}$$

And,

$$\begin{aligned} A^{\frac{1}{n}} &= (\cos\theta + i\sin\theta)^{\frac{1}{n}} \\ &= \cos\left(\frac{2k\pi + \theta}{n}\right) + i\sin\left(\frac{2k\pi + \theta}{n}\right) \end{aligned}$$

Where,

$$k = 0, 1, 2, 3, \dots$$

And,

$$\det\left(A^{\frac{1}{n}}\right) \neq 0$$

Therefore, the mappings  $\phi$  are differentiable homomorphisms from the circle group  $SO_2$  to  $GL_n$ .

## Section 6

1. a

Consider,

The bracket operation,

$$[A, B] = AB - BA.$$

And,

The Jacobi identity,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

Since,

$$\begin{aligned}
& [A, [B, C]] + [B, [C, A]] + [C, [A, B]] \\
&= A[B, C] - [B, C]A + B[C, A] - [C, A]B + C[A, B] - [A, B]C \\
&= A(BC - CB) - (BC - CB)A + B(CA - AC) - (CA - AC)B + C(AB - BA) - (AB - BA)C \\
&= ABC - ACB - BCA + CBA + BCA - BAC - CAB + ACB + CAB - CBA - ABC + BAC \\
&= ABC - ABC - ACB + ACB - BCA + BCA + CBA - CBA - BAC + BAC - CAB + CAB \\
&= 0
\end{aligned}$$

Therefore,

For bracket operation,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

Hence, the bracket operation  $[A, B] = AB - BA$  satisfies the Jacobi identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

## 2. a

Consider,  $V$  be a real vector space of dimension 2 with a law of composition  $[v, w]$  that is bilinear and skew-symmetric.

And,

The Jacobi identity,

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

Since,  $V$  be a real vector space of dimension 2.

Therefore,  $u$  can be written as a linear span of  $v$  and  $w$ .

$$u = c_1 v + c_2 w.$$

Now,

With the help of bilinearity and skew-symmetry,

$$\begin{aligned}
& [u, [v, w]] + [v, [w, u]] + [w, [u, v]] \\
&= [(c_1 v + c_2 w), [v, w]] + [v, [w, (c_1 v + c_2 w)]] + [w, [(c_1 v + c_2 w), v]] \\
&= [c_1 v, [v, w]] + [c_2 w, [v, w]] + [v, [w, c_1 v]] + [v, [w, c_2 w]] + [w, [c_1 v, v]] + [w, [c_2 w, v]] \\
&= c_1 [v, [v, w]] + c_2 [w, [v, w]] + [v, c_1 [w, v]] + [v, c_2 [w, w]] + [w, c_1 [v, v]] + [w, c_2 [w, v]] \\
&= c_1 [v, [v, w]] + c_2 [w, [v, w]] + c_1 [v, [w, v]] + c_2 [v, [w, w]] + c_1 [w, [v, v]] + c_2 [w, [w, v]] \\
&= c_1 [v, [v, w]] + c_2 [w, [v, w]] + c_1 [v, -[v, w]] + c_2 [v, 0] + c_1 [w, 0] + c_2 [w, -[v, w]] \\
&= c_1 [v, [v, w]] + c_2 [w, [v, w]] - c_1 [v, [v, w]] + c_2 [v, (w - w)] + c_1 [w, (v - v)] - c_2 [w, [v, w]] \\
&= (c_1 [v, [v, w]] - c_1 [v, [v, w]]) + (c_2 [w, [v, w]] - c_2 [w, [v, w]]) + c_2 ([v, w] - [v, w]) \\
&+ c_1 ([w, v] - [w, v])
\end{aligned}$$

Therefore,

After cancelling each term on right hand side,

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

Hence, the real vector space  $V$  of dimension 2 with a law of composition  $[v, w]$  satisfies the Jacobi identity  $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ .

## 3. a

Consider the group  $SL_2$  operates by conjugation on the space of trace-zero matrices.

Comment

Step 2 of 2 ^

Since,

There are three elements in the basis of trace-zero matrices:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Let,

$$A \in SL_2(\mathbb{R})$$

Then,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Where,

$$ad - bc = 1$$

As,

The diagonal matrix commutes with every matrix.

So, the space will also be decomposed into three orbits formed by  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ .

Therefore, **the space can be decomposed into three orbits.**

4. a

Consider  $G$  be the group of invertible real matrices of the form  $\begin{bmatrix} a & b \\ 0 & a^2 \end{bmatrix}$ .

Since,

$$e^{\text{trace} A} = \det e^A$$

Then,

$$e^{a+a^2} = \det \begin{bmatrix} e^a & e^b \\ 1 & e^{a^2} \end{bmatrix}$$

$$e^{a+a^2} = e^{a+a^2} - e^b$$

$$e^b = 0$$

So,

$$b = 1$$

Therefore, the Lie algebra be the group of invertible real matrices of the form  $\begin{bmatrix} a & 1 \\ 0 & a^2 \end{bmatrix}$ ,

Now,

The bracket:

$$\begin{aligned} AB - BA &= \begin{bmatrix} a & 1 \\ 0 & a^2 \end{bmatrix} \begin{bmatrix} b & 1 \\ 0 & b^2 \end{bmatrix} - \begin{bmatrix} b & 1 \\ 0 & b^2 \end{bmatrix} \begin{bmatrix} a & 1 \\ 0 & a^2 \end{bmatrix} \\ &= \begin{bmatrix} ab & a+b^2 \\ 0 & a^2b^2 \end{bmatrix} - \begin{bmatrix} ab & a^2+b \\ 0 & a^2b^2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & a-a^2+b^2-b \\ 0 & 0 \end{bmatrix}_{2 \times 2} \end{aligned}$$

Hence, **the Lie algebra will be the group of invertible real matrices of the form  $\begin{bmatrix} a & 1 \\ 0 & a^2 \end{bmatrix}$  and**

**the bracket is  $\begin{bmatrix} 0 & a-a^2+b^2-b \\ 0 & 0 \end{bmatrix}_{2 \times 2}$ .**

## 5. a

Consider the set defined by  $xy = 1$  and the group of invertible diagonal  $2 \times 2$  matrices.

Assume,

$$xy = 1$$

$$zw = 1$$

And,

$$\begin{bmatrix} x & \\ & y \end{bmatrix}, \begin{bmatrix} z & \\ & w \end{bmatrix}$$

Also,

$$\begin{bmatrix} x & \\ & y \end{bmatrix} \begin{bmatrix} z & \\ & w \end{bmatrix} = \begin{bmatrix} xz & \\ & yw \end{bmatrix}$$

Since,

$$xy = 1, zw = 1$$

So,

$$(xy)(zw) = 1$$

$$(xz)(yw) = 1$$

Therefore, the group  $xy = 1$  is subgroup of the group of invertible diagonal  $2 \times 2$  matrices.

Now,

Its Lie algebra,

$$e^x e^y = e^1$$

$$e^{x+y} = e^1$$

$$x + y = 1$$

Hence, the group  $xy = 1$  is subgroup of the group of invertible diagonal  $2 \times 2$  matrices and its Lie algebra is  $x + y = 1$ .

## 6. a

(a)

To show that  $O_2$  operates by conjugation on its lie algebra;

Since it is known that the determinant of orthogonal matrix  $2 \times 2$  is either 1 or  $-1$ , so the orthogonal group  $O_2$  has group action on the plane that is a rotation.

$$O_2 = \left\{ A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \text{ or } \begin{bmatrix} -\cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}; \det A = \pm 1 \right\}$$

It is also known that the lie algebra of the orthogonal group  $O_2$  consists of the skew-symmetric matrices.

Consider the one parameter group  $\phi(t) = e^{At}$ , where  $A$  is skew symmetric matrix.

Substitute  $t = 0$  in  $\phi(t) = e^{At}$  and  $\frac{d\phi(t)}{dt} = Ae^{At}$

$$\begin{aligned} \phi(t) &= e^{A(0)} \\ &= I \end{aligned}$$

And,

$$\begin{aligned} \frac{d\phi(0)}{dt} &= Ae^{A(0)} \\ &= A \end{aligned}$$

Thus, each matrix  $A$  is skew symmetric is a tangent vector to  $O_2$  at the identity and element of its lie algebra.

Hence,  $O_2$  operates by conjugation on its lie algebra.



(b)

To show that this operation is compatible with the bilinear form  $\langle A, B \rangle = \frac{1}{2} \text{trace} AB$  ;

Since one parameter group  $\phi(t) = e^{At}$  and  $\phi(t) = e^{Bt}$  where  $A$  is skew symmetric matrix.

Differentiate the parameter group with respect to  $t$  and substitute  $t = 0$ , to get

$$\frac{d\phi(t)}{dt} = Ae^{At}$$

$$\begin{aligned} \frac{d\phi(0)}{dt} &= Ae^{A(0)} \\ &= A \end{aligned}$$

And,

$$\frac{d\phi(t)}{dt} = Be^{Bt}$$

$$\begin{aligned} \frac{d\phi(0)}{dt} &= Be^{B(0)} \\ &= B \end{aligned}$$

The operation  $\frac{d\phi(0)}{dt} = A$  and  $\frac{d\phi(0)}{dt} = B$  can be written as,

$$\langle A, B \rangle = \frac{1}{2} \text{trace} AB$$

Hence, operation is compatible with the bilinear form  $\langle A, B \rangle = \frac{1}{2} \text{trace} AB$  .

(c)

To define a homomorphism  $O_2 \rightarrow O_2$  and describe this homomorphism explicitly;

Define  $\psi : O_2 \rightarrow O_2$  be a homomorphism by,

$$\psi(A) = B$$

Where,  $A, B \in O_2$

And apply the operation  $\frac{d\phi(0)}{dt} = A$  and  $\frac{d\phi(0)}{dt} = B$  then the homomorphism can be defined

by,

$$\boxed{\psi\left(\frac{d\phi(0)}{dt}\right) = \frac{d\phi(0)}{dt}}$$

This is differentiable that is homomorphism and  $\frac{d\phi(0)}{dt}$  is the derivative of  $\phi(t)$ , where  $\phi(t)$  is one parameter group.

7. a

Consider the Lie algebras of  $U_n$ ,  $SU_n$ ,  $O_{3,1}$  and  $SO_n(\mathbb{C})$ .

[Comment](#)

Step 2 of 5 ^

(a)

Consider the special unitary group  $U_n$ .

If  $\varphi$  is a path in  $U_n$  with  $\varphi(0) = I$  and  $\dot{\varphi}(0) = A$ , then  $\varphi(t)^* \varphi(t) = I$  identically,

And so,

$$\frac{d}{dt}(\varphi(t)^* \varphi(t)) = 0$$

Then,

$$\begin{aligned} \frac{d}{dt}(\varphi^* \varphi)_{t=0} &= \left( \frac{d\varphi^*}{dt} \varphi + \varphi^* \frac{d\varphi}{dt} \right)_{t=0} \\ &= A^* + A \\ &= 0 \end{aligned}$$

Hence, every skew-hermitian  $n \times n$  matrix is a tangent vector to  $U_n$  at the identity – an element of its lie algebra.

(b)

Consider the special unitary group  $SU_n$ .

If  $\varphi$  is a path in  $SU_n$  with  $\varphi(0) = I$  and  $\dot{\varphi}(0) = A$ , then  $\varphi(t)^* \varphi(t) = I$  identically,

And,

$$\det(\varphi(t)) = 1$$

Therefore,

$$\frac{d}{dt} \det(\varphi(t)) = 0$$

Evaluating at  $t = 0$ ,

$$\text{trace}(A) = 0$$

Also,

$$\frac{d}{dt}(\varphi(t)^* \varphi(t)) = 0$$

Then,

$$\begin{aligned} \frac{d}{dt}(\varphi^* \varphi)_{t=0} &= \left( \frac{d\varphi^*}{dt} \varphi + \varphi^* \frac{d\varphi}{dt} \right)_{t=0} \\ &= A^* + A \\ &= 0 \end{aligned}$$

Hence, every skew-hermitian trace-zero  $n \times n$  matrix is a tangent vector to  $SU_n$  at the identity – an element of its lie algebra.

(c)

Consider the special unitary group  $O_{3,1}$ .

If  $\varphi$  is a path in  $O_{3,1}$  with  $\varphi(0) = I$  and  $\dot{\varphi}(0) = A$ , then  $\varphi' I_{3,1} \varphi = I_{3,1}$  identically,

And so,

$$\frac{d}{dt}(\varphi' I_{3,1} \varphi) = 0$$

Then,

$$\begin{aligned} \frac{d}{dt}(\varphi' I_{3,1} \varphi) &= \left( \frac{d\varphi'}{dt} I_{3,1} \varphi + \varphi' I_{3,1} \frac{d\varphi}{dt} \right)_{t=0} \\ &= A' I_{3,1} + I_{3,1} A \\ &= 0 \end{aligned}$$

Hence, every  $4 \times 4$  matrix  $A$  such that  $A' I_{3,1} = -I_{3,1} A$  is a tangent vector to  $O_{3,1}$  at the identity – an element of its lie algebra.

(d)

Consider the special unitary group  $SO_n(\mathbb{C})$ .

If  $\varphi$  is a path in  $SO_n(\mathbb{C})$  with  $\varphi(0) = I$  and  $\dot{\varphi}(0) = A$ , then  $\varphi' \varphi = I$  identically,

And,

$$\det(\varphi(t)) = 1$$

Therefore,

$$\frac{d}{dt} \det(\varphi(t)) = 0$$

Evaluating at  $t = 0$ ,

$$\text{trace}(A) = 0$$

Also,

$$\frac{d}{dt}(\varphi' \varphi) = 0$$

Then,

$$\begin{aligned} \frac{d}{dt}(\varphi' \varphi)_{t=0} &= \left( \frac{d\varphi'}{dt} \varphi + \varphi' \frac{d\varphi}{dt} \right)_{t=0} \\ &= A' + A \\ &= 0 \end{aligned}$$

Hence, every trace-zero  $n \times n$  matrix such that  $A' = -A$  is a tangent vector to  $SO_n(\mathbb{C})$  at the identity – an element of its lie algebra.

8. a

Consider the Lie algebra of  $SP_{2n}$  using block form  $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ .

---

Consider the special unitary group  $SP_{2n}$ .

If  $\varphi$  is a path in  $SP_{2n}$  with  $\varphi(0) = I$  and  $\varphi'(0) = M$ , then  $\varphi' S \varphi = S$  identically,

Where,

$$S = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

And,

$$\det(\varphi(t)) = 1$$

Therefore,

$$\frac{d}{dt} \det(\varphi(t)) = 0$$

Evaluating at  $t = 0$ ,

$$\text{trace}(A + D) = 0$$

Also,

$$\frac{d}{dt}(\varphi' S \varphi) = 0$$

Then,

$$\begin{aligned} \frac{d}{dt}(\varphi' S \varphi) &= \left( \frac{d\varphi'}{dt} S \varphi + \varphi' S \frac{d\varphi}{dt} \right)_{t=0} \\ &= M' S + S M \\ &= 0 \end{aligned}$$

Thus,

$$M' S = -S M$$

$$\begin{aligned} \left[ \begin{array}{c|c} A & C \\ \hline B & D \end{array} \right] \left[ \begin{array}{c|c} 0 & I \\ \hline -I & 0 \end{array} \right] &= - \left[ \begin{array}{c|c} 0 & I \\ \hline -I & 0 \end{array} \right] \left[ \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \\ \left[ \begin{array}{c|c} -C & A \\ \hline -D & B \end{array} \right] &= \left[ \begin{array}{c|c} -C & -D \\ \hline A & B \end{array} \right] \end{aligned}$$

As a result,

$$A + D = 0$$

So obviously,

$$\text{trace}(A + D) = 0$$

Hence, every  $2n \times 2n$  block matrix  $\left[ \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$  such that  $A + D = 0$  is a tangent vector to  $SP_{2n}$  at the identity – an element of its lie algebra.

9. a

Consider, the vector cross product in  $\mathbb{R}^3$ ,  $SU_2$  and  $SO_3$ .

Comment

Step 2 of 3 ^

(a)

Let,

$$F = \mathbb{R}$$

And,

$$L = \mathbb{R}^3$$

Also,

Define  $[\cdot, \cdot]: \mathbb{R}^3 \rightarrow \mathbb{R}^3$

The cross product of vectors:  $[x, y] = x \wedge y$

Now,

The vector cross product satisfies Jacobi identity.

Therefore,

The vector cross product makes  $\mathbb{R}^3$  into Lie algebra  $L_1$ .

Hence, **the vector cross product makes  $\mathbb{R}^3$  into Lie algebra  $L_1$ .**

(b)

Let,

$$L_2 = Lie(SU_2)$$

And,

$$L_3 = Lie(SO_3)$$

As,

The Lie group isomorphism from  $SU_2$  to  $SO_3$ :

$$\phi: SU_2 \rightarrow SO_3$$

Such that:

$$\phi\left(\exp\left[-\sum_k t^k i \frac{\sigma_k}{2}\right]\right) = \exp\left[-\sum_k t^k i l_k\right]$$

The conjugacy classes in  $SO_3$  operate on  $\mathbb{R}^3$  as rotations.

And,

The conjugation by an element of  $SU_2$  rotate every latitude,

Also,

The spin homomorphism  $SU_2 \rightarrow SO_3$  can be used to relate the conjugacy classes in the two groups.

Hence, **the three Lie algebras  $L_1, L_2$  and  $L_3$  are isomorphic.**

10. a

### Classification of Complex Lie Algebras

Let  $L$  be denote a finite dimensional Lie algebra over a field  $F$ .

Note that the notation  $ad_x$ , for  $x \in L$  is given by the map from  $L$  to  $L$  assigning

$$ad_x(y) = [y, x], \text{ for all } y \in L.$$

Now the classical form of  $L$  is defined as the map

$$\langle, \rangle : L \times L \rightarrow F$$

by the assignment

$$\langle x, y \rangle = Tr(ad_x \cdot ad_y).$$

Let us now classify the Lie algebras of dimension one.

Clearly we have

$$L = Fx \text{ for some } x \in L \text{ where } [x, x] = 0.$$

This follows that for all  $y$  and  $z$  in  $L$

$$y = ax \text{ and } z = bx \implies [y, z] = [ax, bx] = ab[x, x] = 0.$$

Hence  $L$  is an abelian and clearly unique up to isomorphism.

This follows that  $L = F^{(-)}$  is the only Complex Lie Algebra of dimension one (upto isomorphism).

Let us now classify the Complex Lie Algebras of dimension two.

Now we have

$$L = Fx + Fy, \text{ for some linearly independent } x, y \in L.$$

In the above argument the  $x, y$  satisfying

$$[x, x] = [y, y] = 0.$$

Therefore only the product  $[x, y]$  which needs to be considered are

- (1) If  $[x, y] = 0$  then  $L$  is abelian.
- (2) If  $[x, y] \neq 0$  then define  $z = ax + by$ , where  $a, b \in F$  are not both zero.

Without loss of generality let us assumed that  $a \neq 0$ .

Then note that

$$[w, z] = z \text{ where } w = a^{-1}y.$$

So it concludes that

$$L = Fw + Fz$$

is a Lie algebra satisfying

$$L' = Fz.$$

By the above argument it follows that this is the only non-abelian two-dimensional Lie algebra up to isomorphism.

And there exists another Lie algebra of type (b) above, name as  $L_2$ .



Let us now classify the Complex Lie Algebras of dimension three.

Let us consider a basis of  $L$  as  $(x_1, x_2, x_3)$ . Since we have  $[L, L] = L$ ,

$$(f_1, f_2, f_3) := ([x_2, x_3], [x_3, x_1], [x_1, x_2])$$

is also basis of  $L$ .

Hence notice that  $f_i$  as linear combinations of the  $x_i$ , that is,

$$f_i = \sum a_{ij} e_j, \text{ with non-singular coefficient matrix } A = (a_{ij}).$$

Now notice that  $A$  must be symmetric, and that every invertible symmetric matrix  $A$  determines a Lie algebra  $L_A$ .

Then two such Lie algebras  $L_A$  and  $L_B$  are isomorphic if and only if there is a matrix  $M \in GL_3(\mathbb{C})$  such that

$$B = \det(M)(M^{-1})^t A M^{-1}.$$

In other words,  $L_A$  and  $L_B$  are isomorphic if and only if  $A$  and  $B$  lie in the same  $G$ -orbit under the action of  $G := \mathbb{C}^* \times GL_3(\mathbb{C})$  given by

$$(t, C)A = tCAC^T \text{ on the space of symmetric } 3 \times 3 \text{ matrices.}$$

Recall the Principal axis theorem, the system of representatives is just the identity. Hence there is only one complex Lie algebra of dimension 3 (upto isomorphism).

## Result

4 of 4

First we classify the dimension one complex Lie algebras  $L$  upto isomorphism then of dimensions two and three consecutively.

11. a

Consider  $B$  be a real  $n \times n$  matrix and  $\langle \cdot, \cdot \rangle$  be the bilinear form  $X'BY$ .

And,

The orthogonal group  $G$  of this form is defined to be the group of matrices  $P$  such that:

$$P'BP = B.$$

Since,

$$e^{\text{trace} A} = \det e^A$$

Let,

$$A \in G$$

So,

$$A'B = BA^{-1}$$

Thus,

$$A'S + SA = 0$$

Therefore,

The Lie algebra is:

$$L = \{A \in GL_n(\mathbb{R}) : A'S + SA = 0\}$$

Also,

The one parameter group is:

$$G = \{A \in GL_n(\mathbb{R}) : e^{At}, A'S + SA = 0\}$$

Hence, the Lie algebra is  $L = \{A \in GL_n(\mathbb{R}) : A'S + SA = 0\}$  and one parameter group will be

$$G = \{A \in GL_n(\mathbb{R}) : e^{At}, A'S + SA = 0\}.$$

## Section 7

1. a

**To Prove:** The Unitary group  $U_n$  is path connected.

**Solution:** Let us start with a statement of Spectral theorem.

Spectral theorem states that if  $U$  is an  $n \times n$  unitary matrix then there exists an another unitary matrix  $P$  and also  $P$  is invertible such that  $PUP^{-1}$  is a diagonal matrix.

We will propose to prove that the group  $U_n$  is path connected.

In order to prove this, it is enough to show that there exists a continuous map joining  $PUP^{-1}$  and the identity matrix,  $I$ .

Let us now consider the diagonal matrix  $PUP^{-1}$  by the way

$$PUP^{-1} = \text{diag}(\exp(i\alpha_1), \exp(i\alpha_2), \dots, \exp(i\alpha_n)), \quad \text{where } \alpha_i \in \mathbb{R}.$$

Let  $f(x)$  be the continuous curve in the diagonal unitary matrices such that the diagonal entries of  $f(x)$  are  $\exp(ix\alpha_j)$ , where  $1 \leq j \leq n$  and  $x \in [0, 1]$ .

This follows that

$$f(0) = I \quad \text{and} \quad f(1) = PUP^{-1}.$$

Let us now consider a function  $g$  from  $[0, 1]$  defined by the assignment

$$g(x) = P^{-1}f(x)P.$$

Then  $f$  is a continuous curve in  $U_n$  such that

$$g(0) = I \quad \text{and} \quad g(1) = U.$$

So we construct a continuous curve  $g$  in  $U_n$  such that

$$g(0) = I \quad \text{and} \quad g(1) = U.$$

Hence  $U_n$  is path connected.

This completes the proof.

---

### Result

2 of 2

Choose an arbitrary element from  $U_n$  and shown that there exists a continuous path in  $U_n$  joining the element and the identity matrix.

2. a

To determine the dimensions of the following groups:

[Comment](#)

Step 2 of 8 ^

(a)

Consider the following group:

$$U_n$$

Since it is known that it is a unitary matrix of subgroup of  $GL_n(\mathbb{C})$ .

Suppose that  $u \in U_n$  then different columns of  $u$  are orthogonal, that is

$$\sum_i u_{ij}^* u_{ik} = 0 \dots\dots (1)$$

Where,  $j \neq k$

And each column of  $u$  is normalized to unity,

$$\sum_i u_{ij}^* u_{ij} = 1 \dots\dots (2)$$

Where,  $j = k$

Now to find the dimension of  $U_n$ ;

From equation (1), for the first row, a complex constraint equation is  $n-1$ .

For the second row, a complex constraint equation is  $n-2$  and so on.

This implies that, number of complex equation is  $\frac{n(n-1)}{2}$  or real equation is  $n(n-1)$ .

From equation (2), number of real equations is  $n$ .

So, dimension of  $U_n$  is:

$$\begin{aligned} d &= 2n^2 - (n(n-1) + n) \\ &= 2n^2 - (n^2 - n + n) \\ &= n^2 \end{aligned}$$

Hence, dimension of  $U_n$  is  $n^2$ .

(b)

Consider the following group:

$$SU_n$$

Since the special unitary group in  $n$  dimensions and it is known that the  $SU_n$  is subgroup of  $U_n$ .

Determinant of every matrix of  $SU_n$  is 1.

Suppose that  $u \in U_n$  then determinant of matrices of  $U_n$  is  $e^{i\alpha}$  where  $\alpha$  is real number.

So, number of equation is real matrix of determinant of  $u$  is 1 this implies that determinant of  $u$  is 1.

Hence, dimension of  $SU_n$  is  $n^2 - 1$ .

(c)

Consider the following group:

$$SO_n(\mathbb{C})$$

Since it is known that  $SO_n(\mathbb{C})$  is subgroup of  $O_n(\mathbb{C})$  and  $O_n$ .

Suppose  $f \in O_n(\mathbb{C})$  this means that  $f^t f = 1$  where  $f$  is  $n \times n$  matrix and  $\det F = \pm 1$

Now consider the part of  $O_n(\mathbb{C})$  with  $\det F = +1$

This is known as  $SO_n(\mathbb{R}) \equiv SO(n)$  and dimension of  $SO_n(\mathbb{R})$  is  $\frac{n(n-1)}{2}$ .

Hence, dimension of  $SO_n(\mathbb{C})$  is  $\boxed{n^2 - n}$

(d)

Consider the following group:

$$O(3,1)$$

First consider the group  $O(p,q)$  and suppose that  $f \in O(p,q)$  then

A matrix  $A$  is invariant under  $O(p,q)$ , this means that

$$f^t A f = A$$

It is the real subgroup of  $O(p,q)$ .

Here,  $p = 3$  and  $q = 1$  then  $O(3,1)$  is known as Lorentz group and dimension is  $\frac{n(n-1)}{2}$ .

Hence, dimension of  $O(3,1)$  is  $\boxed{\frac{n(n-1)}{2}}$ .

(e)

Consider the following group:

$$SP_{2n}$$

This group is known as symplectic group.

Suppose that matrix  $A \in SP_{2n}(\mathbb{C})$  can be defined as,

$$A^t M A = M$$

[Comment](#)

Step 8 of 8 ^

Where, matrix  $M$  invariant in  $SP_{2n}$ :

$$M = \begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ -I_{n \times n} & 0_{n \times n} \end{bmatrix}$$

Suppose  $A \in SP_{2n}(\mathbb{R})$  then it is also a real.

Hence, dimension of  $SP_{2n}$  is  $n(2n-1) = \boxed{2n^2 - n}$ .

3. a

**Solution:** We will find all solutions of the equation  $P^2 = I$  near  $I$ .

Let us recall that for some  $A$  we have

$$e^A = I + \sum_k \frac{A^k}{k!}.$$

Now consider the exponential map  $\phi$  defined by the assignment

$$\phi(X) = e^X.$$

Then  $\phi$  is a local diffeomorphism around  $X = 0$ .

Then there exists a positive  $\epsilon$  such that  $\phi$  is a diffeomorphism in the neighbourhood  $N_\epsilon(0)$  of  $X = 0$ .

Let us now choose  $\delta > 0$  in such a way that

$$N_\delta(e^0) = N_\delta(I) \subseteq \phi(N_{\frac{\epsilon}{2}}(0)).$$

Therefore we have if  $\|P - I\| < \delta$  then there exists a unique  $X$  such that

$$\|X\| < \frac{\epsilon}{2} \quad \text{and} \quad e^X = P.$$

Let us now assume that the aforementioned  $P$  satisfied

$$\|P - I\| < \delta \quad \text{and} \quad P^2 = I.$$

Let us again assume that  $X$  be the unique one for which

$$e^X = P \quad \text{and} \quad \|X\| < \frac{\epsilon}{2}.$$

Then we have

$$\begin{aligned} P^2 &= (e^X)^2 \\ &= e^{2X} \\ &= I = e^0. \end{aligned}$$

Now notice that since  $\phi$  is a diffeomorphism in a neighbourhood  $N_\epsilon(0)$  of  $X = 0$  and  $\|2X\| \leq \epsilon$  we have

$$2X = 0.$$

This follows that

$$X = 0 \quad \text{and} \quad P = I.$$

Therefore the only solution of the equation  $P^2 = I$  in the neighbourhood  $N_\delta(I)$  is  $P = I$ .

This completes the solution.

## Result

Considering a map  $\phi$  by  $\phi(A) = e^A$  and showed that the only solution of the equation  $P^2 = I$  in a neighbourhood of  $I$  is  $P = I$ .

4. a

Consider a path-connected, non-abelian subgroup of  $GL_2$  of dimension 2.

[Comment](#)

Step 2 of 2 ^

Consider, the group  $GL_2$  of dimension 2 containing  $2 \times 2$  matrices such as  $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ ,

Where,

$$a^2 + b^2 \neq 0$$

And,

The group  $SO_2$  is a subgroup of  $GL_2$ .

Also,

The group  $SO_2$  has  $2 \times 2$  matrices such as  $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ ,

Where,

$$a^2 + b^2 = 1$$

As,

It is isomorphic to circle group.

So,

It is path-connected.

Clearly,

It is non-abelian.

Hence, group  $SO_2$  is a path-connected, non-abelian subgroup of  $GL_2$  of dimension 2.

## 5. a

Hermitian matrix is defined on any square matrix that is, the conjugate is equals to the given matrix itself.

**a.**

Consider the map;

$$e: H(n) \rightarrow H(n)$$

Here,  $H$  denotes an hermitian matrix and which is bijective.

For the case of complex invertible matrices the polar form which states that the non-zero complex number;

$$z = a + ib$$

This is given by;

$$z = r(\cos \theta + i \sin \theta)$$

So, by the application of polar coordinates it can be said that there is a bijection between the topological space that is,  $GL(n, \mathbb{C})$  having invertible matrices of order  $n \times n$

Also,  $GL(n, \mathbb{C})$  is a group

Thus, from above information, for every complex invertible matrix  $X$  there exists a unique unitary matrix  $U$  and unique hermitian matrix  $H$  such that;

$$X = Ue^{iH}$$

**Therefore, there is a bijection between  $GL(n, \mathbb{C})$  and  $H(n)$**



b.

From the above part there is an exponential relation with hermitian matrix defined as;

$$X = Ue^H$$

Since,  $U$  be some unitary matrix, so;

$$|\det U| = 1$$

As,  $U$  is considered to be unitary, so from above and  $\text{tr}(H)$  is real and  $H$  being hermitian, so;

$$\det(e^H) > 0$$

Thus, if;

$$\det(X) = 1$$

So, this must obtain;

$$\det(e^H) = 1$$

This implies that;

$$H \in H(n)$$

Thus, there is a bijection between  $SL(n, \mathbb{C})$  and  $H(n) \times H(n)$

**Therefore, the topological structure or the map of  $GL_2(\mathbb{C})$  using the polar decomposition will be defined as**  $SL(n, \mathbb{C}) \rightarrow H(n) \times H(n)$

6. a

Tangent vector field is defined as the generalized vector field on the manifold with the derived tangent expression.

[Comment](#)

Step 2 of 2 ^

Consider the tangent vector field  $PA$  in  $\mathbb{C}$

Given that;

$$A = 1 + i$$

Suppose;

$$P = (x, y)$$

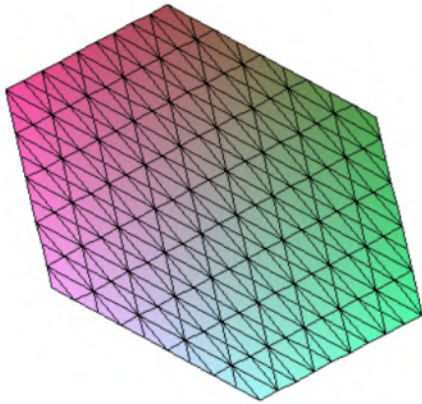
Then, the dot product will be;

$$\begin{aligned} PA &= (x, y) \cdot (1 + i) \\ &= x + iy \end{aligned}$$

Now,  $\mathbb{C}$  will be the group of complex number under multiplication

Take the whole plane as the group  $\mathbb{C}$

So, appropriate graph for the tangent  $PA$  on  $\mathbb{C}$  has been shown below using MAPLE;



7. a

Consider the subgroup  $H$  be a finite normal subgroup of a path connected group  $G$ .

[Comment](#)

Step 2 of 2 ^

Since,

The subgroup  $H$  is a finite normal subgroup of a path connected group  $G$ .

As,

The center of a path connected group  $G$  is also path connected and normal subgroup.

And it is connected to all the members of the path connected group  $G$ .

So,

The center of a path connected group  $G$  is maximal path connected normal subgroup.

And,

The subgroup  $H$  be a finite normal subgroup of a path connected group  $G$ ,

Therefore,

If the subgroup  $H$  is a finite normal subgroup of a path connected group  $G$  then  $H$  will be contained in the center of  $G$ .

Because,

The center of a path connected group  $G$  is maximal path connected normal subgroup.

Hence, **the subgroup  $H$  is contained in the center of  $G$ .**

## Section 8

1. a

Groups are defined as an algebraic structure consisting of a set of elements combined with operations closure, associativity, identity and invertibility.

[Comment](#)

## Step 2 of 4 ^

Let  $F$  be a field of order at least four

This implies that only proper normal subgroup of  $SL_2(F)$  is its centre  $Z = \{\pm 1\}$

Claim: projective group  $PSL_2(F)$  is a simple group

Now, using the result that only proper normal subgroup of  $SL_2(F)$  is its centre  $Z = \{\pm 1\}$  and using the correspondence theorem which states that;

Let  $\phi: G \rightarrow \mathcal{G}$  be a surjective group homomorphism with kernel  $K$ . There is a bijective correspondence between subgroups of  $\mathcal{G}$  and subgroup of  $G$  that contain  $K$ , that is;

$$\{\text{subgroups of } G \text{ that contain } K\} \leftrightarrow \{\text{subgroups of } \mathcal{G}\}$$

Using the above result of correspondence theorem the projective groups  $PSL_2(F)$  will be a simple finite group where  $F$  is a finite field.

Also, there is a result that the order of a finite field is always a power of a prime, that for every prime power  $q$ ;

$$q = p^e$$

Here, the field is of the form of  $\mathbb{F}_4$ . So;

$$\begin{aligned} q &= 4 \\ &= 2^2 \end{aligned}$$

This implies that;

$$\begin{aligned} q &= 4 \\ p &= 2 \end{aligned}$$

Where,  $p$  is a prime

And, also  $\mathbb{F}_4$  has a characteristic 2. It is known that finite fields of order  $2^n$  have characteristic 2. That is in those fields the fields will be of the form of;

$$\begin{aligned} 1 &= -1 \\ I &= -I \end{aligned}$$

**Thus, for a field  $F = \mathbb{F}_4$  the projective group  $PSL_2(F)$  is a simple group with its centre as  $Z = \{\pm I\}$**

Next, the field is of the form of  $\mathbb{F}_5$ . So;

$$\begin{aligned} q &= 5 \\ &= 5 \times 1 \end{aligned}$$

This implies that;

$$\begin{aligned} q &= 5 \\ p &= 1, 5 \end{aligned}$$

Where,  $p$  is a prime

And, also  $\mathbb{F}_5$  has a characteristic 2. It is known that finite fields of order odd multiples have characteristic 2. That is in those fields the fields will be of the form of;

$$\begin{aligned} 1 &= -1 \\ I &= -I \end{aligned}$$

**Thus, for a field  $F = \mathbb{F}_5$  the projective group  $PSL_2(F)$  is a simple group with its centre as  $Z = \{\pm I\}$**

2. a

Consider the isomorphism  $PSL_2(F_2) \cong S_3$  and  $PSL_2(F_3) \cong A_4$ .

For  $PSL_2(F_2)$ , consider six elements:

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}$$

And,

The projective line has three elements:

$$\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

So,

An isomorphism can be defined between the permutation group on the projective line elements and  $S_3$ .

Because,

Both are groups of all permutations on three elements.

Therefore,

$$PSL_2(F_2) \cong S_3$$

Similarly,

Every matrix in  $PSL_2(F_3)$  can be decomposed into an even number of transpositions on the four elements.

Thus,

$$PSL_2(F_3) \cong A_4$$

Hence, the isomorphism  $PSL_2(F_2) \cong S_3$  and  $PSL_2(F_3) \cong A_4$ .

### 3. a

A Sylow  $p$ -subgroup of any group  $G$  is defined as the maximal  $p$ -subgroup of  $G$  that is not a proper subgroup of any other  $p$ -subgroup of  $G$ .

a.

Let  $p$  be some prime.

Then the order of  $p$ -Sylow will be of the form of;

$$|S_p| = p!$$

And, the highest power which divides  $|S_p|$  will be  $p$

Thus,  $p$ -subgroups are precisely the cyclic subgroup of order  $p$ , which is generated by  $p$ -cycle

Thus, the number Sylow  $p$ -subgroup is of the form of;

$$\frac{p!}{p}$$

For,  $p = 2$ , it will be;

$$\begin{aligned} \frac{2!}{2} &= \frac{2}{2} \\ &= 1 \end{aligned}$$

For,  $p = 3$ , it will be;

$$\begin{aligned} \frac{3!}{3} &= \frac{6}{3} \\ &= 2 \end{aligned}$$

For,  $p = 5$ , it will be;

$$\begin{aligned} \frac{5!}{5} &= \frac{120}{5} \\ &= 24 \end{aligned}$$

**Therefore, the number of Sylow  $p$ -subgroups of  $PSL_2(\mathbb{F}_5)$  for  $p = 2, 3, 5$  will be 1, 2, 24 respectively.**

b.

Next, to prove that the three subgroups  $A_5$ ,  $PSL_2(\mathbb{F}_4)$  and  $PSL_2(\mathbb{F}_5)$  are isomorphic

Clearly, the order of  $A_5$  is;

$$\begin{aligned} \frac{5!}{2} &= \frac{120}{2} \\ &= 60 \end{aligned}$$

Also, the order of  $PSL(2, n)$  is of the form of;

$$\frac{(n^2 - 1)}{2}$$

Now, for  $PSL_2(\mathbb{F}_4)$  the order will be;

$$\begin{aligned} \frac{(4^2 - 1)}{2} &= \frac{15}{2} \\ &= \frac{60}{2} \\ &= 30 \end{aligned}$$

And, for  $PSL_2(\mathbb{F}_5)$  the order will be;

$$\begin{aligned} \frac{(5^2 - 1)}{2} &= \frac{24}{2} \\ &= \frac{120}{2} \\ &= 60 \end{aligned}$$

Now, since  $A_5$  is simple unique non-abelian group of smallest order.

This means that  $A_5$  is the simple non-abelian group of smallest order possible and is isomorphic to the simple order as of  $A_5$  and of an divisible order

Since, the projective special linear groups are simple. So, the groups  $PSL_2(\mathbb{F}_4)$  and  $PSL_2(\mathbb{F}_5)$  are simple

Also, 60 and 30 are divisible

**Therefore,  $A_5$ ,  $PSL_2(\mathbb{F}_4)$  and  $PSL_2(\mathbb{F}_5)$  are isomorphic to each other.**

#### 4. a

Symplectic group is defined as the two different but closely related collections in the mathematical groups denoted by  $S_p(2n, F)$

a.

To construct a polynomial equation that defines the symplectic group.

Take an example;

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

This is defined as the Lorentz group which preserves each of the two connected components of the hyperboloid of the form;

$$x_1^2 + x_2^2 + x_3^2 - x_4^2 = -1$$

A similar construction can be done with taking any skew symmetric matrix of order  $n \times n$

Say, this matrix is  $S$ , then;

$$S^T = -S$$

If  $\det S \neq 0$

Then it implies that  $n$  would be even

That is, say;

$$n = 2k$$

Take any example say  $X$ ;

$$X = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

And, so get;

$$X_{2k} = \begin{bmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A \end{bmatrix}$$

The matrix group defined as;

$$\text{Symp}_{2n}(\mathbb{R}) = \{A \in GL_{2k}(\mathbb{R}) : AX^T A = X_{2k}\} \\ \leq GL_{2k}(\mathbb{R})$$

This is called the  $2k \times 2k$  real symplectic group

**Therefore, the polynomial in symplectic group can be of the form of;**

$$x_1^2 + x_2^2 + x_3^2 - x_4^2 = -1$$



b.

Consider;

$$X = \{x_{ij}\} \\ \in M_n(\mathbb{C})$$

Then clearly;

$$X^* = \left(\overline{X}\right)^T \\ = \overline{\left(X^T\right)}$$

This is defined as the Hermitian conjugate of  $X$ , that is;

$$\left(X^*\right)_{ij} = \overline{x_{ji}}$$

Then,  $n \times n$  unitary group is the subgroup defined by;

$$U(n) = \{X \in GL_n(\mathbb{C}) : X^*X = I\} \\ \leq GL_n(\mathbb{C})$$

Now, the unitary condition states to  $n^2$  number of equations for  $n^2$  number of complex number, say  $x_{ij}$

Hence, the unitary group  $U_n$  can be defined by the real polynomial equations .

5. a

Consider the centers of the groups  $SL_n(\mathbb{R})$  and  $SL_n(\mathbb{C})$ .

[Comment](#)

Step 2 of 2 ^

Since,

The multiples of the identity follows commutative property with every element, so in the center.

Thus,

The center of  $SL_n(\mathbb{R})$  is  $\{I, -I\}$ .

And,

The center of  $SL_n(\mathbb{C})$  contains the matrices of the form  $kI$ , where  $k^n = 1$ .

Hence, the centers of  $SL_n(\mathbb{R})$  is  $\{I, -I\}$  and centers of  $SL_n(\mathbb{C})$  includes the matrices of the form  $kI$ , where  $k^n = 1$ .

6. a

Consider the normal subgroups of  $GL_2(\mathbb{R})$ .

[Comment](#)

Step 2 of 2 ^

Since,

The normal subgroups of  $GL_2(\mathbb{R})$  that contain its center are:

The special linear group  $SL_2(\mathbb{R})$ ,

The special orthogonal group  $SO_2(\mathbb{R})$ ,

And,

The group  $SU_2(\mathbb{R})$

Therefore, **the normal subgroups of  $GL_2(\mathbb{R})$  that contain its center are  $SL_2(\mathbb{R})$ ,  $SO_2(\mathbb{R})$  and  $SU_2(\mathbb{R})$ .**

7. a

Consider, the isomorphism between the groups  $PSL_n(\mathbb{C})$  and  $GL_n(\mathbb{C})/Z$  as well as  $PSL_n(\mathbb{R})$  and  $GL_n(\mathbb{R})/Z$ .

[Comment](#)

Step 2 of 2 ^

Since,

The groups  $PSL_n(\mathbb{C})$  and  $GL_n(\mathbb{C})/Z$  are not isomorphic because there is no isomorphism between  $PSL_n(\mathbb{C})$  and  $GL_n(\mathbb{C})/Z$ .

Now,

Consider the isomorphism  $\phi$  between  $PSL_n(\mathbb{R})$  and  $GL_n(\mathbb{R})/Z$ .

Then,

$$\text{Ker}(\phi) = \pm I$$

This is one coset of  $\{\pm I\}$  in  $GL_n(\mathbb{R})$  and one member of  $PSL_n(\mathbb{R})$ .

And, center  $Z$  of  $GL_n(\mathbb{R})$  is  $\{\pm I\}$ .

Therefore, **the groups  $PSL_n(\mathbb{C})$  and  $GL_n(\mathbb{C})/Z$  are not isomorphic while the groups  $PSL_n(\mathbb{R})$  and  $GL_n(\mathbb{R})/Z$  are isomorphic.**

8. a

Matrix is defined as the representation of elements using the rows and columns with related orders.

[Comment](#)

Step 2 of 4 ^

a.

Consider  $P$  to be a matrix in the centre of  $SO_n$

Then  $P$  has a property, that it will remain unchanged even after applying the transpose property of a matrix

To show: that  $PA = AP$

Since, given that  $A$  is skew symmetric, then it can be defined as;

$$A' = -A$$

Now, consider  $PA$ , and then take transpose of this;

$$\begin{aligned}(PA)' &= A'P' \\ &= -AP'\end{aligned}$$

Since, the matrix in  $SO_n$  are commutative, then;

$$\begin{aligned}-AP' &= -P'A \\ &= P'(-A) \\ &= (AP')'\end{aligned}$$

Therefore, it is clear that  $AP = PA$

b.

Use the method of mathematical induction for the proof. First consider without loss of generality;

$$q = qn$$

Where,  $q$  is in centre of  $SO_n$

Now, for any  $n > 1$

There will be smooth affine surface defined as;

$$H = \{q = 1\}$$

Clearly, this is irreducible.

Then by using Witt's extension theorem which states that;

Let  $(X, a)$  be a finite-dimensional vector space over an arbitrary field  $K$  along with the nondegenerate symmetric or skew-symmetric bilinear form. If  $f: U \rightarrow U'$  is an isometry between two subspaces of  $X$  then  $f$  extends to an isometry of  $X$

Then by using the result of above stated theorem there exist a finite dimensional quadratic space over a field  $K$

This space acts transitively on the set  $K$  embedding into a field quadratic.

Now, for  $n > 1$ , take  $n = 2$

Then the matrix will be of the form of say,  $M^n$ , where for  $n = 1$  the matrix is defined below;

$$M^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

So, clearly from the above explanation the assumption for  $n \geq 4$  is true

First taking,  $n = 2$ ;

$$\begin{aligned} M^2 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= I \end{aligned}$$

Again for  $n = 3$ ;

$$\begin{aligned} M^3 &= M^2 \times M \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Next for  $n = 4$ ;

$$\begin{aligned} M^4 &= M^3 \times M \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= I \end{aligned}$$

Finally for  $n = 5$ ;

$$\begin{aligned} M^5 &= M^4 \times M \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Therefore it is proved that the centre of  $SO_n$  is trivial if  $n$  is odd and is  $\{\pm I\}$  if  $n$  is even for  $n \geq 4$

## 9. a

Find the order of the general case:

$$SO_n(\mathbb{F}_q)$$

To compute the orders of the groups;

Consider the homomorphism  $\det : GL_n(F) \rightarrow F^\times$ . This map is surjective.

Whole space of  $F^\times$  is the image of  $GL_n(F)$ .

Consider the  $n \times n$  matrix:

$$\begin{pmatrix} x & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

The above matrix is an invertible  $n \times n$  matrix of determinant  $x$ .

It is known that  $SO_n(\mathbb{F}_q)$  is kernel of the homomorphism, then from the first isomorphism theorem,

$$\frac{GL_n(\mathbb{F}_q)}{SL_n(\mathbb{F}_q)} \cong F^\times$$

Therefore,

$$\begin{aligned} |SO_n(\mathbb{F}_q)| &= \frac{|GL_n(\mathbb{F}_q)|}{|F^\times|} \\ |SO_n(\mathbb{F}_q)| &= \frac{(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{n-1})}{(q - 1)} \quad \dots (1) \end{aligned}$$

(a)

Consider the following group:

$$SO_2(\mathbb{F}_3)$$

Substitute  $n = 2, q = 3$  in equation (1), then

$$\begin{aligned} |SO_2(\mathbb{F}_3)| &= \frac{(3^2 - 3^{2-1})(3^2 - 3^{2-2})}{(3 - 1)} \\ &= \frac{(3^2 - 3^1)(3^2 - 3^0)}{2} \\ &= \frac{(9 - 3)(9 - 1)}{2} \\ &= \frac{6 \times 8}{2} \\ &= 24 \end{aligned}$$

Hence, order of  $SO_2(\mathbb{F}_3)$  is  $\boxed{24}$ .

(b)

Consider the following group:

$$SO_3(\mathbb{F}_3)$$

Substitute  $n = 3, q = 3$  in equation (1), then

$$\begin{aligned} |SO_3(\mathbb{F}_3)| &= \frac{(3^3 - 3^{3-1})(3^3 - 3^{3-2})(3^3 - 3^{3-3})}{(3 - 1)} \\ &= \frac{(3^3 - 3^2)(3^3 - 3^1)(3^3 - 3^0)}{2} \\ &= \frac{(27 - 9)(27 - 3)(27 - 1)}{2} \\ &= \frac{18 \times 24 \times 26}{2} \\ &= 5616 \end{aligned}$$

Hence, order of  $SO_3(\mathbb{F}_3)$  is  $\boxed{5616}$ .

(c)

Consider the following group:

$$SO_2(\mathbb{F}_5)$$

Substitute  $n = 2, q = 5$  in equation (1), then

$$\begin{aligned} |SO_2(\mathbb{F}_5)| &= \frac{(5^2 - 5^{2-1})(5^2 - 5^{2-2})}{(5-1)} \\ &= \frac{(25-5)(25-1)}{4} \\ &= \frac{20 \times 24}{4} \\ &= 120 \end{aligned}$$

Hence, order of  $SO_2(\mathbb{F}_5)$  is  $\boxed{120}$ .

(d)

Consider the following group:

$$SO_3(\mathbb{F}_5)$$

Substitute  $n = 3, q = 5$  in equation (1), then

$$\begin{aligned} |SO_3(\mathbb{F}_5)| &= \frac{(5^3 - 5^{3-1})(5^3 - 5^{3-2})(5^3 - 5^{3-3})}{(5-1)} \\ &= \frac{(5^3 - 5^2)(5^3 - 5^1)(5^3 - 5^0)}{4} \\ &= \frac{(125 - 25)(125 - 5)(125 - 1)}{4} \\ &= \frac{100 \times 120 \times 124}{4} \\ &= 372000 \end{aligned}$$

Hence, order of  $SO_3(\mathbb{F}_5)$  is  $\boxed{372000}$ .

10. a

(a)

To write the matrix of conjugation by  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  on  $V$  in block form;

Let  $V$  be the space  $V$  of complex  $2 \times 2$  matrix, with the basis  $(e_{11}, e_{12}, e_{21}, e_{22})$ .

$$e_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$e_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$e_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$e_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$



Then, conjugation by  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  on  $V$  as shown below,

$$\begin{aligned} T(e_{11}) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \\ &= ae_{11} + 0 \cdot e_{12} + ce_{21} + 0 \cdot e_{22} \end{aligned}$$

$$\begin{aligned} T(e_{12}) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \\ &= 0 \cdot e_{11} + ae_{12} + 0 \cdot e_{21} + ce_{22} \end{aligned}$$

$$\begin{aligned} T(e_{21}) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} b & 0 \\ d & 0 \end{bmatrix} \\ &= be_{11} + 0 \cdot e_{12} + de_{21} + 0 \cdot e_{22} \end{aligned}$$

And,

$$\begin{aligned} T(e_{22}) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix} \\ &= 0 \cdot e_{11} + be_{12} + 0 \cdot e_{21} + de_{22} \end{aligned}$$

The coefficients of the  $e_{ij}$ 's on the right is the columns of our matrix. Put it all together, the desired matrix is:

**Hence**, matrix of conjugation by  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  on  $V$  in block forms as shown below:

$$\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$$

(b)

To prove that conjugation defines a homomorphism  $\varphi: SL_2(\mathbb{C}) \rightarrow GL_4(\mathbb{C})$  and that the image of  $\varphi$  is isomorphic to  $PSL_2(\mathbb{C})$ .

Define a map  $\varphi: SL_2(\mathbb{C}) \rightarrow GL_4(\mathbb{C})$  by,

$$\varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$$

To show that  $\varphi$  is homomorphic.

Let  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} f & g \\ h & i \end{bmatrix} \in SL_2(\mathbb{C})$  then,

$$\begin{aligned} \varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} f & g \\ h & i \end{bmatrix}\right) &= \varphi\left(\begin{bmatrix} a+f & b+g \\ c+h & d+i \end{bmatrix}\right) \\ &= \begin{bmatrix} a+f & 0 & b+g & 0 \\ 0 & a+f & 0 & b+g \\ c+h & 0 & d+i & 0 \\ 0 & c+h & 0 & d+i \end{bmatrix} \\ &= \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix} + \begin{bmatrix} f & 0 & g & 0 \\ 0 & f & 0 & g \\ h & 0 & i & 0 \\ 0 & h & 0 & i \end{bmatrix} \\ &= \varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) + \varphi\left(\begin{bmatrix} f & g \\ h & i \end{bmatrix}\right) \end{aligned}$$

And,

$$\begin{aligned} \varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} f & g \\ h & i \end{bmatrix}\right) &= \varphi\left(\begin{bmatrix} af+bh & ag+bi \\ cf+dh & cg+di \end{bmatrix}\right) \\ &= \begin{bmatrix} af+bh & 0 & ag+bi & 0 \\ 0 & af+bh & 0 & ag+bi \\ cf+dh & 0 & cg+di & 0 \\ 0 & cf+dh & 0 & cg+di \end{bmatrix} \\ &= \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix} \begin{bmatrix} f & 0 & g & 0 \\ 0 & f & 0 & g \\ h & 0 & i & 0 \\ 0 & h & 0 & i \end{bmatrix} \\ &= \varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \varphi\left(\begin{bmatrix} f & g \\ h & i \end{bmatrix}\right) \end{aligned}$$

Hence, map  $\varphi: SL_2(\mathbb{C}) \rightarrow GL_4(\mathbb{C})$  homomorphic.

Define a map  $\phi: SL_2(\mathbb{C}) \rightarrow PSL_2(\mathbb{C})$  by,

$$\phi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}{\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}}$$

Where,  $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{C})$  and  $\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$  is the subgroup of scalar transformations with unit determinant.

Since  $\varphi$  is homomorphism from  $SL_2(\mathbb{C})$  to  $GL_4(\mathbb{C})$  then from the isomorphic theorem:

$$PSL_2(\mathbb{C}) \cong \frac{SL_2(\mathbb{C})}{SZ_2(\mathbb{C})}$$

Hence, image of  $\varphi$  is isomorphic to  $PSL_2(\mathbb{C})$ .

(c)

To prove that  $PSL_2(\mathbb{C})$  is a complex algebraic group by finding polynomial equations in the entries  $y_{ij}$  if a  $4 \times 4$  matrix whose solutions are the matrices in the image of  $\phi$ ;

Consider the following  $4 \times 4$  matrix:

$$A = \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$$

Now substitute  $y_{ij}$  in the above entries, then the matrix become,

$$A = \begin{bmatrix} y_{11} & 0 & y_{13} & 0 \\ 0 & y_{22} & 0 & y_{24} \\ y_{13} & 0 & y_{33} & 0 \\ 0 & y_{24} & 0 & y_{44} \end{bmatrix}$$

This can be written as polynomial equations in the entries  $y_{ij}$  as shown below:

$$y(x) = (y_{11} + y_{13}) + (y_{22} + y_{24})x + (y_{13} + y_{33})x^2 + (y_{24} + y_{44})x^3$$

Or,

$$y(x) = (y_{24} + y_{44})x^3 + (y_{13} + y_{33})x^2 + (y_{22} + y_{24})x + (y_{11} + y_{13})$$

**Hence,**  $PSL_2(\mathbb{C})$  is a complex algebraic group.

## Miscellaneous Problem

1. a

Let  $G = SL_2(\mathbb{R})$  and  $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$  be a matrix in  $G$ .

Suppose that  $t$  be its trace.

Substitute  $w = t - x$  in matrix  $A$ ,

$$A = \begin{bmatrix} x & y \\ z & t - x \end{bmatrix}$$

Determinant of  $A$  is  $\det A = x(t - x) - yz$  and it is given that  $\det A = 1$ , so

$$x(t - x) - yz = 1$$

This implies that,

$$xt - x^2 - yz = 1$$

This implies that,

$$xt = x^2 + yz + 1$$

This implies that,

$$t = \frac{x^2 + yz + 1}{x}$$

So, here  $x, y$  and  $z$  are in  $\mathbb{R}$  and fixed trace  $t$  is in  $x, y$  and  $z$  terms.

Decompose them into conjugacy classes:

Suppose that matrix  $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$  is in subgroup  $G$ .

The conjugacy classes in  $G$  can be defined as,

$$\begin{aligned} \text{CL}(A) &= \{BXB^{-1} : B \in G\} \\ BXB^{-1} &= \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & -bx+y \\ z & -bz+w \end{pmatrix} \\ &= \begin{pmatrix} x+bz & -bx+y-b^2z+bw \\ z & -bz+w \end{pmatrix} \end{aligned}$$

Suppose that  $x+bz=a, -bx+y-b^2z+bw=c, -bz+w=e$

And determinant of the above matrix is invertible. So it is form a conjugacy class of  $G$

Thus, conjugacy classes in  $G$  is  $\left\{ \begin{pmatrix} a & c \\ z & e \end{pmatrix} ; c \in \mathbb{R} \right\}$ .

2. a

Consider the elements of  $SL_2(\mathbb{R})$  lie on a one-parameter group.

[Comment](#)

Step 2 of 2 ^

Since,

The one-parameter groups in  $SL_2(\mathbb{R})$  are the homomorphisms  $t \rightarrow e^{tA}$ , where  $A$  is a real  $2 \times 2$  matrix, whose trace is zero.

Hence, **the elements of  $SL_2(\mathbb{R})$  lie on a one-parameter group consists of the form  $e^{tA}$ , where  $A$  is a real  $2 \times 2$  matrix, whose trace is zero.**

3. a

Consider the conjugacy classes in a path connected group  $G$ .

[Comment](#)

Step 2 of 2 ^

Since,

The group  $G$  is a path connected group.

As,

The conjugacy classes in a path connected group can be considered as the equivalence classes of the free loops under the free homotopy.

And,

A free loop in  $G$  is the equivalence class of continuous functions from the circle to  $G$ .

Also,

The two loops are equivalent if they have difference by reparametrization of the circle.

Therefore, it can be viewed as a string.

In a string, all the points between two given points situated on that string only because string is path connected.

Therefore, **the conjugacy classes in a path connected group are path connected.**

4. a

Consider,

$$\alpha = a + bi + cj + dk.$$

Where  $a, b, c$  and  $d$  are real numbers.

[Comment](#)

Step 2 of 5 ^

(a)

Since,

$$\bar{\alpha} = a - bi - cj - dk.$$

Now,

$$\begin{aligned}\bar{\alpha}\alpha &= (a - bi - cj - dk)(a + bi + cj + dk) \\ &= a^2 - b^2(ii) - c^2(jj) - d^2(kk) + (ab - ba)i + (ac - ca)j + (ad - da)k \\ &\quad + [-bc(ij) - cb(ji)] + [-bd(ik) - db(ki)] + [-cd(jk) - dc(kj)]\end{aligned}$$

Then,

$$\begin{aligned}\bar{\alpha}\alpha &= a^2 - b^2(-1) - c^2(-1) - d^2(-1) + (0)i + (0)j + (0)k \\ &\quad + [-bc(k) - cb(-k)] + [-bd(-j) - db(j)] + [-cd(i) - dc(-i)]\end{aligned}$$

So,

$$\bar{\alpha}\alpha = a^2 + b^2 + c^2 + d^2.$$

Therefore,  $\boxed{\bar{\alpha}\alpha = a^2 + b^2 + c^2 + d^2}$ .

(b)

Consider,

$$\alpha \neq 0.$$

Then,

$$a^2 + b^2 + c^2 + d^2 \neq 0.$$

Assume,

$$\alpha^* = \frac{\bar{\alpha}}{a^2 + b^2 + c^2 + d^2}.$$

Since,

$$\begin{aligned}\alpha^* \alpha &= \left( \frac{\bar{\alpha}}{a^2 + b^2 + c^2 + d^2} \right) \alpha \\ &= \frac{\bar{\alpha} \alpha}{a^2 + b^2 + c^2 + d^2} \\ &= \frac{a^2 + b^2 + c^2 + d^2}{a^2 + b^2 + c^2 + d^2} \\ &= 1\end{aligned}$$

And,

$$\begin{aligned}\alpha \alpha^* &= \alpha \left( \frac{\bar{\alpha}}{a^2 + b^2 + c^2 + d^2} \right) \\ &= \frac{\alpha \bar{\alpha}}{a^2 + b^2 + c^2 + d^2} \\ &= \frac{a^2 + b^2 + c^2 + d^2}{a^2 + b^2 + c^2 + d^2} \\ &= 1\end{aligned}$$

As,

$$\alpha \alpha^* = \alpha^* \alpha = 1.$$

So,

The multiplicative inverse of  $\alpha$  is  $\alpha^*$ .

Therefore, **every**  $\alpha \neq 0$  have a multiplicative inverse.

(c)

Consider,

The elements of  $SU_2$  are complex  $2 \times 2$  matrices of the form,

$$A = \begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix},$$

Where,

$$\det A = a^2 + b^2 + c^2 + d^2 = 1.$$

Therefore, this defines a bijective correspondence of  $SU_2$  with the unit 3-sphere

$$\{a^2 + b^2 + c^2 + d^2 = 1\} \text{ in } \mathbb{R}^4.$$

And,

The 3-sphere  $\{a^2 + b^2 + c^2 + d^2 = 1\}$  has a group structure.

Now,

$$\begin{bmatrix} a_1 + b_1 i & c_1 + d_1 i \\ -c_1 + d_1 i & a_1 - b_1 i \end{bmatrix} \begin{bmatrix} a_2 + b_2 i & c_2 + d_2 i \\ -c_2 + d_2 i & a_2 - b_2 i \end{bmatrix} = \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix},$$

Where,

$$x_0 = a_1 a_2 - b_1 b_2 - c_1 c_2 - d_1 d_2,$$

$$x_1 = a_1 b_2 + b_1 a_2 + c_1 d_2 - d_1 c_2,$$

$$x_2 = a_1 c_2 + c_1 a_2 - b_1 d_2 + d_1 b_2,$$

$$x_3 = a_1 d_2 + d_1 a_2 + b_1 c_2 - c_1 b_2.$$



Also,

$$(a_1 + b_1i + c_1j + d_1k)(a_2 + b_2i + c_2j + d_2k) = x_0 + x_1i + x_2j + x_3k.$$

As,

$$(x_0 + x_1i + x_2j + x_3k) \leftrightarrow \begin{bmatrix} x_0 + x_1i & x_2 + x_3i \\ -x_2 + x_3i & x_0 - x_1i \end{bmatrix},$$

So, the map from quaternions  $a^2 + b^2 + c^2 + d^2 = 1$  to  $SU_2$  that sends  $a + bi + cj + dk$  to

$$\begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix} \text{ is an isomorphism.}$$

Hence, **the set of quaternions  $\alpha$  such that  $a^2 + b^2 + c^2 + d^2 = 1$  forms a group under multiplication that is isomorphic to  $SU_2$ .**

5. a

Consider the affine group  $A_n$  is the group of transformations of  $\mathbb{R}^n$  generated by  $GL_n$  and the group  $T_n$  of translations  $t_a(x) = x + a$ .

Since,

The affine group  $A_n$  is the group of transformations of  $\mathbb{R}^n$  generated by  $GL_n$ .

And,

The group  $T_n$  of translations  $t_a(x) = x + a$

For  $A \in A_n$  and  $T \in T_n$ ,

$$\begin{aligned} ATA^{-1}(x) &= AT[A^{-1}(x)] \\ &= A[A^{-1}(x) + a] \\ &= A[A^{-1}(x)] + Aa \\ &= AA^{-1}(x) + Aa \end{aligned}$$

$$= I(x) + Aa$$

$$= x + Aa$$

So,

$$ATA^{-1} \in T_n$$

Therefore, the group  $T_n$  is normal subgroup of  $A_n$ .

Now,

If  $A_n$  is considered as  $n$ -manifold then  $T_n$  will be  $1$ -manifold because  $t_a(x) = x + a$ , so the kernel of the mapping  $\varphi: A_n \rightarrow GL_n$  will be  $T_n$  as  $T_n$  is  $1$ -manifold.

Thus,

By first isomorphism theorem,  $A_n/T_n$  will be isomorphic to  $GL_n$ .

Hence, **the group  $T_n$  is normal subgroup of  $A_n$  and  $A_n/T_n$  is isomorphic to  $GL_n$ .**

6. a

Let  $U$  denote the set of matrices  $A$  such that  $I + A$  is invertible, and define

$$A' = (I - A)(I + A)^{-1}$$

[Comment](#)

Step 2 of 6 ^

(a)

To prove that if  $A$  is in  $U$  then  $A'$  in  $U$ .

Since  $A$  can be written as  $A = (I + A) - I$  and  $I + A$  is invertible then,  $A$  is invertible

And  $I - A = (I + A) - 2A$  then  $I - A$  is invertible.

Thus,  $(I - A)(I + A)^{-1}$

To show that  $I + A'$  is invertible.

$$I + A' = I + (I - A)(I + A)^{-1}$$

Here,  $(I - A)(I + A)^{-1}$  is invertible.

Thus,  $I + A'$  is invertible.

**Hence,**  $A'$  in  $U$

(b)

Prove that the rule  $A \mapsto (I - A)(I + A)^{-1}$  defines a homomorphism from neighborhood of 0 in  $V$  to a neighborhood of  $I$  in  $SO_n$ .

Let  $V$  denote the vector space of real skew-symmetric  $n \times n$  matrices.

Define a map  $A: V \rightarrow SO_n$  by,

$$A(B) = (I - A)B$$

Where,  $B$  is skew-symmetric  $n \times n$  matrix and it is known that determinant in  $SO_n$  is 1.

To show that the map is homomorphism;

Let  $B$  and  $C$  in  $V$

$$\begin{aligned} A(B+C) &= (I - A)(B+C) \\ &= (I - A)B + (I - A)C \\ &= (I - (I - A)(I + A)^{-1})B + (I - (I - A)(I + A)^{-1})C \end{aligned}$$

Since  $(I - A)(I + A)^{-1}$  is invertible LHS part provided in the above is irreducible.

Thus,  $A(B+C) = A(B) + A(C)$

**Hence,** rule  $A \mapsto (I - A)(I + A)^{-1}$  defines a homomorphism from neighborhood of 0 in  $V$  to a neighborhood of  $I$  in  $SO_n$ .

(c)

Yes, there is an analogous statement for the unitary group because  $I + A$  is invertible and matrix  $A$  in  $U$  is symplectic that is determinant is 1 then this set of matrix forms a unitary group.

(d)

Suppose that a matrix  $A$  in  $U$  is symplectic.

To show that  $(A')^t S = -SA'$

Suppose  $S = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$  and a matrix  $A$  in  $U$  is symplectic, that is  $A$  has determinant 1.

$$\begin{aligned} (A')^t S &= ((I-A)(I+A)^{-1})^t \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \\ &= (I+A)^{-t} (I-A)^t \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & (I+A)^{-t} (I-A)^t \\ -(I+A)^{-t} (I-A)^t & 0 \end{bmatrix} \\ &= -(I-A)(I+A)^{-1} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \\ &= -A'S \\ &= -SA' \end{aligned}$$

Hence,  $(A')^t S = -SA'$

Suppose that  $(A')^t S = -SA'$ .

To show that a matrix  $A$  in  $U$  is symplectic, that is to show that  $A$  has determinant 1.

$$\begin{aligned} (A')^t S &= -SA' \\ (A')^t \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} &= -\begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} A' \\ \begin{bmatrix} 0 & (A')^t \\ -(A')^t & 0 \end{bmatrix} &= -\begin{bmatrix} 0 & A' \\ -A' & 0 \end{bmatrix} \end{aligned}$$

This implies that,

$$\begin{bmatrix} 0 & (A')^t + A' \\ -(A')^t + A' & 0 \end{bmatrix} = 0$$

This matrix has determinant 1.

Hence, a matrix  $A$  in  $U$  is symplectic.

7. a

!!!

8. a

!!!

9. a

A group is defined as the algebraic structure consisting of set of elements which is combined with operations to form a third element.

Consider the alternating group  $A_5$  having index 2 of symmetric group;

$$G_2 = S_5$$

Now, suppose the mapping be given as;

$$\varphi: SU_2 \rightarrow SO_3$$

Here,  $G_1$  be the inverse image of the icosahedral group

Let  $H$  be any subgroup of  $G$  which is generated by the rotations

And, let  $F$  be the normal subgroup of  $G$  generated by reflections and rotations

Now, fix;

$$H_i = \alpha_i(H)$$

$$F_i = \alpha_i(F)$$

Then,  $H_i$  is a rotation group, and both are considered to be the normal subgroup of  $G_i$

Now, suppose the possible triplets be  $(G_i, H_i, F_i)$

Here,  $G_i$  is generated by  $H_i$  and  $F_i$

Now, further by using the result that every irreducible rotation group occurs upto conjugation and every irreducible reflection rotation group either appears in conjugation or it contains reflection

So, by above result consider the reflections to be  $f_1, \dots, f_n$  where its corresponding fixed points are defined as the hyper planes known as the walls of a chamber of a reflection group  $W$ .

That is,  $(W, F)$  is a system of reflections with;

$$F = \{f_1, \dots, f_n\}$$

Also, it is given that  $G_3$  is the inverse image of icosahedral group in  $SU_2$

**Therefore, any groups of the form of  $G_i$  is isomorphic to  $H_i$  or  $F_i$**

10. a

!!!

11. a

!!!

12. a

!!!

13. a

Consider the adjoint representation of  $SL_2(\mathbb{C})$ .

[Comment](#)

Step 2 of 2 ^

Since,

The adjoint representation of  $SL_2(\mathbb{C})$  is the representation by conjugation on  $SO_3(\mathbb{C})$ ,

$$\varphi: SL_2(\mathbb{C}) \times SO_3(\mathbb{C}) \rightarrow SO_3(\mathbb{C})$$

As,

$$P \in SL_2(\mathbb{C}), A \in SO_3(\mathbb{C}) \rightarrow PAP^{-1} \in SO_3(\mathbb{C}).$$

The kernel of  $\varphi$  is the center of  $SL_2(\mathbb{C})$  that is  $\{\pm I\}$ .

So,

By first isomorphism theorem,

$$SL_2(\mathbb{C})/\{\pm I\} \approx SO_3(\mathbb{C})$$

Therefore,  $\boxed{SL_2(\mathbb{C})/\{\pm I\} \approx SO_3(\mathbb{C})}$ .



# 10

## Chapter 10

### Section 1

1. a

**To Prove:** The image of a representation of dimension one of a finite group is a cyclic.

**Proof:** Consider the group  $G$  with all non-zero complex numbers, that is

$$G := \mathbb{C} - \{0\}.$$

Then the image of a representation of dimension one of a finite group is a finite subgroup of  $G$ .

So to prove the content, it suffices to show that every finite subgroup of  $G$  is cyclic.

Let  $z \in G$ . Then  $z \neq 0$  and  $z \in \mathbb{C}$ .

Then  $z$  can be written in the form

$$z = r(\cos \theta + i \sin \theta), \text{ where } r > 0 \text{ and } 0 \leq \theta < 2\pi.$$

Let us consider  $w$  be another element with the representation

$$w = R(\cos \phi + i \sin \phi).$$

Then notice that

$$zw = rR(\cos(\theta + \phi) + i \sin(\theta + \phi)).$$

Let us now assume  $H$  is a finite subgroup of  $G$  and  $z \in H$ . Since  $H$  is finite, then there exists a natural number  $n$  such that

$$z^n = 1.$$

This follows that

$$r^n(\cos(n\theta) + i \sin(n\theta)) = 1.$$

But by the above equation notice that if  $r^n = 1$  then clearly  $r = 1$  and  $z$  is on the unit circle.

We also have

$$\begin{aligned} n\theta &= 2k\pi, \text{ where } k \in \mathbb{Z} \\ \implies \theta &= \frac{2k\pi}{n}. \end{aligned}$$

So it is left to prove that  $H$  is cyclic.



Let  $\psi$  be the minimal angle among all  $z \in H$ . And consider

$$z = \cos \psi + i \sin \psi.$$

If  $z$  does not generate  $H$ , then consider  $w = \cos \delta + i \sin \delta$  such that

$$w \neq z^k \text{ for all } k \in \mathbb{Z}.$$

This follows from the polar representation of  $z$  and  $w$  that  $\delta$  is not an integral multiple of  $\psi$ .

Since  $\psi$  is minimal, there exists an integer  $m > 0$  such that

$$\psi > \delta - m\psi > 0.$$

It follows from the previous argument that

$$wz^{-m} = \cos(\delta - m\psi) + i \sin(\delta - m\psi) \in H.$$

But this contradicts the minimality of  $\psi$  among all arguments in  $H$ .

Thus our assumption is wrong, hence  $w$  is an integral multiple of  $z$ . Hence  $z$  generates the group  $H$ .

This proves that  $H$  is cyclic.

Hence every finite subgroup of  $G$  is cyclic. The image of a representation of dimension one of a finite group being a finite subgroup of  $G$ , it is cyclic.

This completes the proof.

## Result

3 of 3

The image of a representation of dimension one of a finite group being a finite subgroup of  $\mathbb{C} - \{0\}$  is a cyclic group.

## 2. a

(a)

The octahedral group  $O$  consists of rotational symmetries of a cube. These rotations are generated by three  $90^\circ$  rotations.

First choose a standard basis for  $\mathbb{R}^3$ :

$$\mathbf{B} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

Thus, to write a representation of  $O$ , use the appropriate 2 matrices, then to get

$$A_x = \begin{bmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A_y = \begin{bmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix}$$

$$A_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix}$$

$$A_y = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

And,

$$A_x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, the above representation is a standard representation of the octahedral group  $O$ .

(b)

Consider the dihedral group:

$$D_n$$

Dihedral group defined by,

$$D_n = \langle x, y \mid x^n = y^2 = (xy)^2 = 1 \rangle$$

First choose a standard basis for  $\mathbb{R}^3$ :

$$\mathbf{B} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

Relative to the standard basis of  $\mathbb{R}^3$ , a linear map has the matrix of the form:

$$\rho(r) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Where,  $r$  act through reflection across the  $x$ -axis.

Suppose  $c$  act on  $\mathbb{R}^3$  through counterclockwise rotation by angle  $\frac{2\pi}{n}$  and relative to the standard basis of  $\mathbb{R}^3$ , a linear map has the matrix of the form:

$$\rho(c) = \begin{bmatrix} \cos\left(\frac{2\pi}{n}\right) & -\sin\left(\frac{2\pi}{n}\right) & 0 \\ \sin\left(\frac{2\pi}{n}\right) & \cos\left(\frac{2\pi}{n}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\rho(c) = \begin{bmatrix} \cos\left(\frac{2\pi}{n}\right) & 0 & -\sin\left(\frac{2\pi}{n}\right) \\ 0 & 1 & 0 \\ \sin\left(\frac{2\pi}{n}\right) & 0 & \cos\left(\frac{2\pi}{n}\right) \end{bmatrix}$$

And,

$$\rho(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\left(\frac{2\pi}{n}\right) & -\sin\left(\frac{2\pi}{n}\right) \\ 0 & \sin\left(\frac{2\pi}{n}\right) & \cos\left(\frac{2\pi}{n}\right) \end{bmatrix}$$

Thus, the above representation is a standard representation of dihedral group  $D_n$ .

## Section 2

1. a

Show that the standard three-dimensional representation of the tetrahedral group  $T$  is irreducible as a complex representation.

Suppose that axes of coordinate passes through the mid-point of the three edges  $y_1, y_2, y_3$  are generator of the tetrahedral group  $T$ .

This is rotations by  $x$  the rotation by  $\frac{2\pi}{3}$  around the front vertex and by  $\pi$  around the respective edge.

The matrices in  $\mathbb{R}^3$  of these operations are  $R_{y_1}, R_{y_2}, R_{y_3}$  and  $R_x$  as shown below;

$$R_{y_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$R_{y_2} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$R_{y_3} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_x = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

If there is an invariant subspace then a representation is reducible. Eigenvectors are invariant subspaces for a single operator. Representation  $e_1, e_2$  and  $e_3$  are standard unit vectors is similar as the eigenvectors of  $R_{y_1}, R_{y_2}, R_{y_3}$ . Here the standard unit vectors are not eigenvector of  $R_x$ . It is know that this is in three-dimensional space, so there is at least one of the invariant subspaces has to be one dimensional.

Thus, the representation of  $T$  is irreducible.

So, the standard three-dimensional representation of the tetrahedral group  $T$  is irreducible as a complex representation.

**Hence proved**

## 2. a

**Solution:** Let us consider the standard two-dimensional representation of the dihedral group  $D_n$ .

We will now find the value of  $n$  for which  $D_n$  is an irreducible complex representation.

Now note that the aforesaid group  $D_n$  is generated by the rotation by angle  $\frac{2\pi}{n}$  and a reflection in a line passing through the origin. Let us named the line as  $L$ .

Now there are two cases arise.

**Case-1:** Let us take  $n = 1$  and  $n = 2$ .

It is easy to notice that the line  $L$  is invariant under the rotation by  $\pi$  and  $2\pi$  and hence it an invariant subspace for the action of  $D_n$ .

And follows that the representation is reducible.

So case-1 is done. Now move to case-2.

**Case-2:** Let us consider  $n \geq 3$ .

We will propose to prove that  $D_n$  is irreducible for  $n \geq 3$ .

Notice that the only invariant subspaces for the reflection in  $L$  are  $L$  and  $L^\perp$ .

And note that for  $n \geq 3$  the matrix of rotation has no real eigenvectors, so  $L$  and  $L^\perp$  are not invariant under rotation.

Therefore there are no one-dimensional invariant subspaces, and hence the representation of  $D_n$  is irreducible. This completes the solution.

## Result

For  $n \geq 3$   $D_n$  is an irreducible complex representation.

### 3. a

**Solution:** Given that the vector space  $V$  which is a representation of  $S_3$ .

(a) We will first show that  $V$  contains a nonzero invariant subspace of dimension at most 2.

By the given condition  $u \in V$  such that  $u \neq 0$ . Let us now consider the vector

$$x = u + (1\ 2)u + (1\ 3)u + (2\ 3)u + (1\ 2\ 3)u + (2\ 1\ 3)u.$$

This follows that

$$gx = x \quad \forall g \in S_3.$$

Let us now consider the following cases:

**Case-1:** Let us consider  $x \neq 0$ .

Then it spans a one-dimensional invariant subspace isomorphic to the trivial representation of  $S_3$ .

**Case-2:** Let us consider  $x = 0$  and consider the vector

$$v = u + (1\ 2\ 3)u + (1\ 3\ 2)u.$$

Then we have

$$(1\ 2\ 3)v = (1\ 3\ 2)v = v$$

and

$$(1\ 2)v = (1\ 3)v = (2\ 3)v = (1\ 2)u + (1\ 3)u + (2\ 3)u = x - v = -v.$$

Now if  $v \neq 0$  then it spans a one-dimensional invariant subspace isomorphic to the sign representation of  $S_3$ .

**Case-3:** Let us consider  $v = 0$  and consider the vector

$$w = u + (1\ 2)u.$$

Now same as above, if  $w \neq 0$  then

$$w + (1\ 2\ 3)w + (1\ 3\ 2)w = x = 0.$$

This shows that  $w$  and  $(1\ 2\ 3)w$  span one or two-dimensional representation of  $S_3$ .

**Case-4:** Let us take  $w = v = x = 0$ . Then we have

$$(1\ 2)u = -u.$$

This shows that  $u$  and  $(1\ 2\ 3)u$  span one or two-dimensional representation of  $S_3$ .

This completes the proof of (a).

Now we will propose to prove that all two-dimensional representations of  $S_3$  are isomorphic.

Let us assume  $V$  is irreducible. Then it follows that it should coincide with the invariant subspace constructed in (a).

In above case—1 and case—2 we identified is with the trivial and sign representation respectively.

If in case—3 we get a one-dimensional subspace, then we have

$$(1\ 2)w = w$$

and

$$\begin{aligned}(1\ 2\ 3)w &= kw \\ \implies kw &= (1\ 2)(1\ 2\ 3)w \\ \implies kw &= (2\ 3)w \\ \implies kw &= (1\ 3\ 2)(1\ 2)w \\ \implies kw &= k^2w.\end{aligned}$$

Therefore we have  $k = 1$  and we get a trivial representation.

If in case—4 we get a one-dimensional subspace, then we have

$$(1\ 2)u = -u$$

and

$$(1\ 2\ 3)u = ku.$$

Similarly it follows that  $k = 1$  and we get a sign representation. So now it is enough to prove that two-dimensional representations in cases 3 and 4 are isomorphic.

This follows from the fact that both are isomorphic to the standard representation of  $D_3$ , where we choose  $w$  to be parallel to the axis of one of the reflections in case-3 and we choose  $u$  to be perpendicular to this axis in case-4.

This completes the proof.

## Result

3 of 3

First show that  $V$  contains a nonzero invariant subspace of dimension at most 2 and then we determine all irreducible representations of  $S_3$ .

## Section 3

1. a

**Given:**  $G$  is a cyclic order of order 3 and given a matrix  $A$  of order 3 as

$$A = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

Now  $A$  defines a matrix representation of  $G$ .

**Solution:** We will use the averaging process to produce a  $G$ -invariant form from the standard Hermitian product  $X^*Y$  on  $\mathbb{C}^2$ .

Let us consider the inner product  $\langle, \rangle$  on  $\mathbb{C}^2$  by the assignment

$$\langle X, Y \rangle = X^*Y.$$

Now notice that

$$\begin{aligned} \langle AX, AY \rangle &= (AX)^*(AY) \\ &= (X^*A^*)(AY). \end{aligned}$$

Now let us define a Hermitian product by the assignment

$$\langle X, Y \rangle_1 = \frac{1}{3}(\langle X, Y \rangle + \langle AX, AY \rangle + \langle A^2X, A^2Y \rangle).$$

So basically observe that the above equation may be expressed as

$$\langle X, Y \rangle_1 = X^*BY, \quad \text{where } B = \frac{1}{3}(I + A^*A + (A^*)^2A^2).$$

Let us now calculate them all.

$$\begin{aligned} A^*A &= \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}. \\ A^2 &= \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}. \end{aligned}$$

Therefore

$$\begin{aligned} (A^*)^2A^2 &= \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

Hence we have

$$\begin{aligned} B &= \frac{1}{3}(I + A^*A + (A^*)^2A^2) \\ &= \frac{1}{3} \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \right) \\ &= \begin{bmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{bmatrix}. \end{aligned}$$

This completes the solution.

## Result

3 of 3

Considering the given hermitian product we have define a new hermitian product by using  $A$  and then find the required result by using some matrix operations.



2. a

**Solution:** Given that  $\rho : G \rightarrow GL(V)$  is a representation of a finite group on a real vector space  $V$ .

(a) First we prove that there exists  $G$ -invariant, positive definite, symmetric form  $\langle \cdot, \cdot \rangle$  on  $V$ .

Without loss of generality let us consider a positive definite, symmetric bilinear form  $[\cdot, \cdot]$  on  $V$ . Let us define the averaged form by considering

$$\langle v, w \rangle = \frac{1}{|G|} \sum_{g \in G} [\rho(g)v, \rho(g)w].$$

The form is symmetric, positive definite and  $G$ -invariant. Now we will show that the form is symmetric.

Now we have

$$\begin{aligned} \langle v, w \rangle &= \frac{1}{|G|} \sum_{g \in G} [\rho(g)v, \rho(g)w] \\ &= \frac{1}{|G|} \sum_{g \in G} [\rho(g)w, \rho(g)v] \\ &= \langle w, v \rangle. \end{aligned}$$

Now we will show that the form is positive definite.

So it is enough to show that  $\langle v, v \rangle > 0$ .

Now notice that  $\sum_{g \in G} [\rho(g)v, \rho(g)v]$  is the sum of positive numbers we have

$$\langle v, v \rangle = \frac{1}{|G|} \sum_{g \in G} [\rho(g)v, \rho(g)v] > 0.$$

Now the verification of the  $G$  invariance follows rearranging the summation, once one notices that for any element  $x$  in  $G$  and the right multiplication by  $x$  gives a permutation of  $G$  given by the assignment

$$\begin{aligned} \langle \rho(x)v, \rho(x)w \rangle &= \frac{1}{|G|} \sum_{g \in G} [\rho(g)\rho(x)v, \rho(g)\rho(x)w] \\ &= \frac{1}{|G|} \sum_{gx \in G} [\rho(gx)v, \rho(gx)w] \\ &= \langle v, w \rangle. \end{aligned}$$

(b) We will now show that  $\rho$  is a direct sum of irreducible representations.

Without loss of generality let us assume that  $\rho$  is not irreducible.

Let us consider a subspace  $W$  of  $V$  such that  $W$  is  $\rho$  invariant.

We will propose to prove that the orthogonal of  $W$  with respect to the form  $\langle \cdot, \cdot \rangle$  defined in (a) also  $\rho$ -invariant.

So in order to show the above, it is enough to show that

$$z \in W^\perp \text{ and } g \in G \implies \rho(g)z \in W^\perp.$$

So we will assert that

$$\langle \rho(g)z, w \rangle = 0 \quad \forall w \in W.$$

Therefore we have

$$\begin{aligned} \langle \rho(g)z, w \rangle &= \langle \rho(g^{-1})\rho(g)z, \rho(g^{-1})w \rangle, \text{ since } \rho \text{ is invariant} \\ &= \langle z, \rho g^{-1}w \rangle, \text{ since } \rho \text{ is a representation} \\ &= 0, \text{ since } \rho(g^{-1})w \in W. \end{aligned}$$

This follows that  $\rho$  splits as a direct sum of two representations  $\rho_1$  and  $\rho_2$ .

Now if we use induction on the dimension of  $V$ , we will get the required result as mentioned above.

**Note:** In our case  $V$  is a vector space of finite dimension.

(c) Now we show that finite subgroup  $G$  of  $GL_n(\mathbb{R})$  is conjugate to a subgroup of  $O_n$ .  
Now  $O_n$  is defined by

$$O_n = \{A \in GL(n, \mathbb{R}) \mid A^t A = I\}.$$

Let's define  $\phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  by the assignment

$$\phi(v, w) = \sum_{g \in G} gv \cdot gw$$

where  $\cdot$  is the usual dot product.

Then note that  $\phi$  is a bilinear symmetric form.

Therefore look that

$$\phi(v, v) = \sum_{g \in G} \|gv\|^2 > 0, \quad \forall v \neq 0$$

and

$$\phi(xv, xw) = \sum_{g \in G} gxv \cdot gxw = \sum_{k \in Gx=G} kv \cdot kw = \phi(v, w)$$

Hence  $\phi$  is a  $G$ -invariant inner product and then  $G$  is a subgroup of a conjugate of  $O_n$ .

This completes the proof.

## Result

4 of 4

First we prove that there exists  $G$ -invariant, positive definite, symmetric form  $\langle \cdot, \cdot \rangle$  on  $V$  then we prove the rest by using this representation.

### 3. a

#### (a)

Let  $G$  be a finite group and  $R : G \rightarrow SL_2(\mathbb{R})$  be a faithful representation of a finite group by real  $2 \times 2$  matrices with determinant 1.

Since every finite subgroup of  $GL_n(\mathbb{R})$  is conjugate to a subgroup of  $O_n$ .

So, there exists a representation  $R(G)$  which is equivalent representation such that

$R(G) \leq O_2$  and it is known that  $R$  is faithful.

Thus,  $G$  is isomorphic to  $R(G)$  and  $R(G) \leq O_2$ .

This means that  $G$  is isomorphic to a finite subgroup of  $O_2$ .

**Hence,**  $G$  is a cyclic group.

[Comment](#)

Step 2 of 5 ^

#### (b)

To determine the finite groups that has faithful real two dimensional representations.

Since it is known that every finite subgroup of  $GL_n(\mathbb{R})$  is conjugate to a subgroup of  $O_n$ ,

Now suppose that  $G$  contained in  $O_2$  that is,  $G \subset O_2$  and also suppose that  $H$  is subgroup of  $G$ .

Thus,  $H = G \cap SO_2$

If  $H = G$  then  $H$  is cyclic and it is done.

If  $H \neq G$  then index of  $H$  is 2 that is  $[G:H] = 2$  and quotient group  $G/H$  contained in  $O_2/S_2$ .

$$G/H \subset O_2/SO_2$$

This consists of reflections.

Thus,  $G$  is **dihedral**.

[Comment](#)

Step 4 of 5 ^

(c)

To determine the finite groups that has faithful real three-dimensional representations with determinant 1.

Since it is known that every finite subgroup of  $GL_3(\mathbb{R})$  is conjugate to a subgroup of  $O_3$ ,

Now suppose that  $G$  contained in  $O_3$  that is,  $G \subset O_3$  and also suppose that  $H$  is subgroup of  $G$ .

$$\text{Thus, } H = G \cap SO(3)$$

If  $H = G$  then  $H$  is cyclic and it is done.

If  $H \neq G$  then index of  $H$  is 2 that is  $[G:H] = 2$  and quotient group  $G/H$  contained in  $O_3/SO_3$ .

$$G/H \subset O_3/SO_3$$

And finite subgroup of  $SO(3)$  isomorphic to  $S_4$  or  $A_4$  and it is cyclic or dihedral group.

Thus,  $G$  are **cyclic or dihedral group and  $S_4$  or  $A_4$** .

4. a

Suppose  $V$  has a skew symmetric bilinear form given by  $B : V \times V \rightarrow V$ . So, we have  $B(u, v) = -B(v, u)$ , and  $B(u, u) = 0$ . Suppose  $\rho$  is a representation of  $G$ , that is,  $\rho : G \rightarrow GL(V)$ . We have to produce a  $G$ -invariant skew symmetric bilinear form on  $V$ . Consider  $B' : V \times V \rightarrow V$ , given by

$$B'(u, v) = \frac{1}{|G|} \sum_{g \in G} B(\rho_g(u), \rho_g(v)) \quad (1)$$

It is clear that  $B'$  is skew-symmetric bilinear form. We just check that it is  $G$ -invariant. To do so, let  $h \in G$ . Now,

$$\begin{aligned} B'(\rho_h(u), \rho_h(v)) &= \frac{1}{|G|} \sum_{g \in G} B(\rho_g(\rho_h(u)), \rho_g(\rho_h(v))) = \frac{1}{|G|} \sum_{g \in G} B(\rho_{gh}(u), \rho_{gh}(v)) = \\ &= \frac{1}{|G|} \sum_{g' \in G} B(\rho_{g'}(u), \rho_{g'}(v)) = B'(u, v). \end{aligned}$$

The penultimate equality is true because as  $g$  runs over  $G$ ,  $gh$  also runs over  $G$ .

Now, we show that the new skew symmetric bilinear form  $B'$  need not be non-degenerate. Suppose  $V$  is a vector space over  $\mathbb{C}$ , with dimension 2. Let  $S = \{v_1, v_2\}$ , be a basis of  $V$ . Define  $B : V \times V \rightarrow \mathbb{C}$ , by defining then on the basis elements, as  $B(v_i, v_i) = 0$ , for  $i = 1, 2$ , and  $B(v_1, v_2) = 1, B(v_2, v_1) = -1$ . Observe that  $B$  is completely determined by these values, and it is clearly checked to be a non-degenerate bilinear form. To prove the non-degeneracy observe that if  $B'$ , be the matrix of  $B$ , with respect to the basis  $S$ , then,

$$B' = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Clear  $B'$  is invertible, and hence  $B$  is non-degenerate. Now, consider the the two-dimensional standard representation of  $S_3$ . We write it explicitly, with  $V$  as the representation space, we define as follows:  $\rho : S_3 \rightarrow GL(V)$ , as

$$\rho_e = id, \rho_{(12)} : V \rightarrow V, v_1 \mapsto -v_1, v_2 \mapsto v_1 + v_2,$$

$$\rho_{(23)} : V \rightarrow V, v_1 \mapsto v_1 + v_2, v_2 \mapsto -v_2$$

$$\rho_{(13)} : V \rightarrow V, v_1 \mapsto -v_2, v_2 \mapsto -v_1$$

$$\rho_{(123)} : V \rightarrow V, v_1 \mapsto v_2, v_2 \mapsto -v_1 - v_2$$

$$\rho_{(132)} : V \rightarrow V, v_1 \mapsto -v_1 - v_2, v_2 \mapsto v_1$$

This completes the description of  $\rho$ . Now, Let  $B_1$  be the  $G$ -invariant skew symmetric bilinear form that we have obtained in the previous part. We claim that  $B_1 = 0$ . This will prove that  $B_1$  is degenerate. Now, observe that automatically  $B_1(v_i, v_i) = 0$ , for  $i = 1, 2$ . Now, we compute  $B_1(v_1, v_2)$ .

To do that as in the formula, we have to compute  $B(\rho_g(v_1), \rho_g(v_2))$ . We do that first:

$$B(\rho_e(v_1), \rho_e(v_2)) = B(v_1, v_2) = 1$$

$$B(\rho_{(12)}(v_1), \rho_{(12)}(v_2)) = B(-v_1, v_1 + v_2) = -1$$

$$B(\rho_{(23)}(v_1), \rho_{(23)}(v_2)) = B(v_1 + v_2, -v_2) = -1$$

$$B(\rho_{(13)}(v_1), \rho_{(13)}(v_2)) = B(-v_2, -v_1) = -1$$

$$B(\rho_{(123)}(v_1), \rho_{(123)}(v_2)) = B(v_2, -v_1 - v_2) = 1$$

$$B(\rho_{(132)}(v_1), \rho_{(132)}(v_2)) = B(-v_1 - v_2, v_1) = 1$$

So,  $B_1(v_1, v_2) = \frac{1}{|G|} \sum_{g \in G} B(\rho_g(v_1), \rho_g(v_2)) = \frac{1}{6}[1 - 1 - 1 - 1 + 1 + 1] = 0$ . Henceforth,  $B_1(v_2, v_1) = 0$ , showing that  $B_1$  is identically 0. Thus corresponding to a non-degenerate skew symmetric bilinear form, the  $G$ -invariant bilinear form so obtained by averaging process, need not be non-degenerate.

## Result

3 of 3

For the first part we use the typical averaging process done in the text for hermitian forms. For the second part we construct a counter example by using the standard representation of  $S_3$ . See the solution for more details.

5. a

$G$  be an abelian group of order  $p$ . Let  $x$  be a generator of  $G$ . Define a representation  $\phi : G \rightarrow GL_2(\mathbb{F}_p)$ , given by

$$\phi(x) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Therefore,

$$\phi(x^k) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

Consider the vector

$$v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Now observe that  $\phi(x^k)v = v$ . Therefore  $v$  is a common eigen-vector of  $\phi(g)$ , for every  $g \in G$ . Therefore  $\langle v \rangle$  is a proper  $G$ -invariant subspace. Therefore, we conclude that  $\phi$  is reducible. Now, we show that  $\phi$  cannot be decomposed as the sum of two irreducible representation. Indeed, if one had found another  $G$ -invariant proper subspace disjoint from the previous one, say  $\langle v_1 \rangle$ , then  $v_1$  would have been a common eigen vector for  $\phi_g$ . But, it is clear to observe that each  $\phi_g$  has eigen value 1, with multiplicity 2, but the eigen space corresponding to 1, is one dimension which is  $\langle v \rangle$ . Therefore, no further common eigen-vector exists and therefore, no further one-dimension  $G$ -invariant subspace is present. Therefore,  $G$  cannot be decomposed into two irreducible representation.

## Result

2 of 2

The idea is to find the common eigen spaces of each  $\phi_g$ , where  $\phi$  is the given representation. See the solution for more details

## Section 4

1. a



We have to find the dimensions of the distinct irreducible representations of  $G$ , where  $G$  is the octahedral group or the quaternion group or the dihedral groups  $D_4, D_5, D_6$ . We will do this one group at a time. The main result we are going to use is as follows:

**Result:** Let  $G$  be a finite group and  $\rho_1, \dots, \rho_k$  are the non-isomorphic irreducible representation of  $G$ , then  $k$  is the number of conjugacy classes of  $G$ , and if  $d_1, d_2, \dots, d_k$  denote the degrees respectively, then  $\sum_{i=1}^k d_i^2 = |G|$ .

With this result in hand, we proceed one by one. Let  $G$  be the octahedral group, but since the octahedral group is isomorphic to  $S_4$ , we can assume  $G = S_4$ . The conjugacy class of  $S_n$  in general is well known, and it is known that two elements are conjugate iff they have the same cycle type, and hence the number of conjugacy classes of  $S_n$  is equal to the partition of  $n$ . Therefore  $G$  has five conjugacy classes. Therefore it has five non-isomorphic representation with degrees say  $d_i (1 \leq i \leq 5)$ . Therefore, we have

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 = 24$$

Now, there are only two representation of  $S_4$ , which are one-dimensional. One is the trivial map, and the other the sign map. See exercise 5.2 of section 5, which has an answer by me. So  $d_1 = d_2 = 1$ . This further simplifies the above equation as

$$d_3^2 + d_4^2 + d_5^2 = 22$$

Now, observe that each of the remaining degrees must be less than 5. Now let one of them is 4, say  $d_3 = 4$ . That means  $d_4^2 + d_5^2 = 6$ , but this equation doesn't have a positive integer solution. Therefore, we conclude that the remaining degrees  $d_3, d_4, d_5$  have to all either 2 or 3. Now all of them cannot be equal to 3 as then  $d_3^2 + d_4^2 + d_5^2 = 27$ , which is a contradiction. Similarly all of them cannot be 2. Also, observe that two of them being 2, and the other being 3, also is not equal 22. The only possibility that matches is one of them is 2, and the other remaining 2 are 3. So, we conclude, without loss of generality, that  $d_3 = 3, d_4 = d_5 = 3$ . Hence, we conclude finally that there are  $d_1 = d_2 = 1, d_3 = 2, d_4 = d_5 = 3$ .

Now, let  $G = Q_8$  be the quaternion group. Now  $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ , such that  $(-1).i = -i, (-1).j = -j, (-1).k = -k$ , and  $ij = k, jk = i, ki = j, ji = -k, kj = -i, ik = -j$ . The conjugacy class of  $Q_8$  is well known and is as follows:

$$\{1\}, \{-1\}, \{\pm i\}, \{\pm j\}, \{\pm k\}$$

. Therefore  $G$  has five conjugacy classes. Therefore it has five non-isomorphic representation with degrees say  $d_i (1 \leq i \leq 5)$ . Therefore, we have

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 = 8$$

Now, it is clear that none of them can be more than 3. So, each  $d_i \geq 2$  for every  $1 \leq i \leq 5$ . Again, observe that two of them cannot be true as then the other remaining three has to be all zero, but degrees are positive integers. Also, all of them cannot be 1, as then it will contradict the above equation. Therefore, one of them is 2, and all others have to be 1. So, without loss of generality, we conclude  $d_1 = d_2 = d_3 = d_4 = 1, d_5 = 2$ .

Now, let  $G = D_4$ . We know  $D_4 = \{1, r, r^2, r^3, s, sr, sr^2, sr^3\}$ . Therefore,  $r^4 = 1 = s^2, rs = sr^{-1}$ . Now, again the conjugacy class is well known and is as follows

$$\{1\}, \{r^2\}, \{r, r^3\}, \{s, sr\}, \{sr^2, sr^3\}$$

We see that there are exactly 5 conjugacy classes, and therefore we get ,

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 = 8$$

Now, we argue exactly as in the case of  $Q_8$ , to conclude that  $d_1 = d_2 = d_3 = d_4 = 1, d_5 = 2$ .



Next, we have  $G = D_5 = \{1, r, r^2, r^3, r^4, s, sr, sr^2, sr^3, sr^4\}$ . Therefore,  $r^5 = 1 = s^2, rs = sr^{-1}$ . Now, again the conjugacy class is well known and is as follows

$$\{1\}, \{r, r^4\}, \{r^2, r^3\}, \{s, sr, sr^2, sr^3, sr^4\}$$

We see that there are exactly 4 conjugacy classes, and therefore we get ,

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 = 10$$

Observe that  $d_i \leq 4$ , for every  $1 \leq i \leq 4$ . Observe that none of them can't be 3, because if they do, then,

$$d_1^2 + d_2^2 + d_3^2 = 1$$

which is not possible, as each of them are positive integers. We conclude that the degrees are only 1 and 2. All of them can't be one, and all of them can't be 2. Therefore, there are 3 choices remaining, one of them is that 3 representations have degree 1, and the other of 2, but squares of this combination add upto 6, which is contradiction. Also, three of them can't be 2 for the same reason. The last remaining choice is two of them are degree 1, and the other two are of 2, and this combination indeed matches the equation. Therefore, we conclude,

$$d_1 = d_2 = 1, d_3 = d_4 = 2$$

Finally, let  $G = D_6 = \{1, r, r^2, r^3, r^4, r^5, s, sr, sr^2, sr^3, sr^4, sr^5\}$ . Therefore,  $r^6 = 1 = s^2, rs = sr^{-1}$ . Now, again the conjugacy class is well known and is as follows

$$\{1\}, \{r^3\}, \{r, r^5\}, \{r^2, r^4\}, \{sr, sr^3, sr^5\}, \{s, sr^2, sr^4\}$$

We see that there are exactly 6 conjugacy classes, and therefore we get ,

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 + d_6^2 = 12$$

Observe that  $d_i \leq 4$ , for every  $1 \leq i \leq 6$ . Observe that none of them can't be 3. We conclude that the degrees are only 1 and 2. All of them can't be one, and all of them can't be 2. There is a unique solution to this equation which is that there are four of them which have degree 1, and the remaining two have degree 2. Therefore, we conclude,

$$d_1 = d_2 = 1 = d_3 = d_4, d_5 = d_6 = 2$$

## Result

5 of 5

If  $G$  is octahedral group,  $G \cong S_4$ , then the degrees of irreducible representations are 1,2,3. For  $G$  the quaternion group, it is 1,2. And finally for each of the dihedral groups, the degrees are 1 or, 2. See the solution for more details

2. a

**Given:**  $G$  is a non-abelian group of order 55.

**Solution:** First we will find the Class equation of  $G$ . Now notice that there exists two elements  $x$  and  $y$  in  $G$  such that

$$G := \langle x, y \mid x^{11} = y^5 = 1, yxy^{-1} = x^3 \rangle.$$

Now recall the **Sylow's theorem** once. Then by Sylow's theorem we have the 11-sylow subgroup  $\langle x \rangle$  of  $G$  is normal and there are 11 conjugate 5-subgroups.

Now notice that centralizer of  $x$ , denoted by  $Z(x)$ , contains  $\langle x \rangle$ . And observe that  $Z(x)$  is not equals to  $G$ , so we have

$$Z(x) = \langle x \rangle.$$

Then the counting formula shows that  $|C(x)| = 5$ . This follows similarly that

$$|C(y)| = 11.$$

Now look at the conjugacy class  $C(x)$ . Then observe that

$$\text{the relation } yxy^{-1} = x^3 \implies x^3 \in C(x).$$

Similar;y we have

$$x^9 \in C(x) \text{ since } x^9 = (yxy^{-1})^3 = yx^3y^{-1}.$$

Continuing this process we have

$$C(x) = \{x, x^3, x^4, x^5, x^9\}.$$

Since  $x^{11} = 1$ , we have the conjugacy class for  $x^2$  as

$$C(x^2) = \{x^2, x^6, x^7, x^8, x^{10}\}.$$

Now notice that there are 11 conjugate 5-subgroups, and any two of them intersects trivially. And each of those 11 subgroups contains an element in the conjugacy class of  $y$ . Now we know from the previous argument that  $|C(y)| = 11$ . Therefore the conjugacy classes of  $y^2, y^3, y^4$  also have order 11 each.

Hence the class equation of  $G$  is given by

$$|G| = 1 + 5 + 5 + 11 + 11 + 11 + 11.$$

Now notice that there are 7 irreducible characters.

Since  $\langle x \rangle$  is a normal subgroup of  $G$ , we can show the existence of the quotient group  $G/\langle x \rangle$ . Since  $G$  has 55 elements and  $\langle x \rangle$  has 11, then the quotient group  $G/\langle x \rangle$  contains 5 elements.

Then the quotient group  $G/\langle x \rangle$  is cyclic, since 5 is prime. Hence generated by an element  $z + \langle x \rangle$ , and  $z \notin \langle x \rangle$ .

Let us now consider the canonical homomorphism as

$$\pi : G \rightarrow G/\langle x \rangle.$$

Now assume any representation of  $G/\langle x \rangle$  by the assignment

$$\bar{\theta} : G/\langle x \rangle \rightarrow GL(V).$$

Then notice that

$$\theta = \bar{\theta} \circ \pi.$$

Since  $G/\langle x \rangle$  is cyclic, its irreducible characters are one-dimensional, and there are five of them. And each irreducible characters corresponding to  $G/\langle x \rangle$  determines the irreducible characters of  $G$ . Hence five irreducible characters out of seven have dimension 1 of  $G$ .

Let  $d_i$  be the dimension of the seven irreducible characters of  $G$ , where  $1 \leq i \leq 7$ .

Now note that

$$|G| = \sum_{i=1}^7 d_i^2.$$

Since

$$d_1 = d_2 = d_3 = d_4 = d_5 = 1$$

the dimension of other two irreducible characters of  $G$  must be five to satisfy the above argument.

So there are 7 irreducible characters of  $G$  in which five have dimension 1 and the remaining two have dimension 5.

This completes the solution.

## Result

3 of 3

Class equation of  $G$  is  $|G| = 55 = 1 + 5 + 5 + 11 + 11 + 11 + 11$  and there are 7 irreducible characters of  $G$  in which five have dimension 1 and the remaining two have dimension 5.

3. a

### Character table for Klein four-group

Let  $V$  denotes the Klein four-group. Now  $V$  is an abelian group of order 4. Therefore there are exactly 4 irreducible representations, all of dimension have 1.

Since every non-zero element of  $V$  has order 2, all values of characters are  $\pm 1$ . Therefore the character table be look like as with characters  $\chi_1, \chi_2, \chi_3$  and  $\chi_4$ :

$$\begin{bmatrix} & 1 & x & y & xy \\ \chi_1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & \pm 1 & \pm 1 & \pm 1 \\ \chi_3 & 1 & \pm 1 & \pm 1 & \pm 1 \\ \chi_4 & 1 & \pm 1 & \pm 1 & \pm 1 \end{bmatrix}$$

Now the only way to make all the rows of this table orthogonal, is by taking

$$\begin{bmatrix} & 1 & x & y & xy \\ \chi_1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & 1 & -1 & -1 \\ \chi_3 & 1 & -1 & 1 & -1 \\ \chi_4 & 1 & -1 & -1 & 1 \end{bmatrix}$$

### Character table for the quaternion group

Let  $Q$  denotes the quaternion group.

We begin by finding the number of irreducible characters of  $Q$  and their degree. To do this, we note that if  $d_\alpha$  is the degree of any irreducible character that  $d_\alpha$  divides 8 and  $d_\alpha = 1, 2, 4, 8$ . Moreover, since

$$d_\alpha^2 \leq 8 \text{ we may conclude that } d_\alpha = 1, 2.$$

Hence the required table is

$$\begin{bmatrix} & 1 & -1 & i & j & k \\ \chi_1 & 1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & 1 & 1 & -1 & -1 \\ \chi_3 & 1 & 1 & -1 & -1 & 1 \\ \chi_4 & 1 & 1 & -1 & 1 & -1 \\ \chi_5 & 2 & -2 & 0 & 0 & 0 \end{bmatrix}$$

#### Character table for the dihedral group $D_4$

Now recall that

$$D_4 = \{x, y \mid x^4 = y^2 = xyxy = 1\}.$$

The Class equation of  $D_4$  is given by

$$|D_4| = 1 + 1 + 1 + 1 + 4.$$

So  $D_4$  can be represented as follows

$$D_4 = \{1, x, x^2, x^3, y, xy, x^2y, x^3y\}.$$

Now from the above mention statement it follows that  $D_4$  has 5 classes. So have 5 characters.

Among the 5 characters, the first 4 are one-dimensional. And we will find them.

The center of  $D_4$  is the set of elements which are alone in their conjugacy class. Therefore we have

$$Z(D_4) = \{1, x^2\}.$$

Hence center of  $D_4$  is a normal subgroup of  $D_4$  with quotient isomorphic to Klein four-group.

Hence the character table for  $D_4$  is

$$\begin{bmatrix} & 1 & -1 & i & j & k \\ \chi_1 & 1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & 1 & -1 & 1 & -1 \\ \chi_3 & 1 & 1 & -1 & -1 & 1 \\ \chi_4 & 1 & 1 & 1 & -1 & -1 \\ \chi_5 & 2 & -2 & 0 & 0 & 0 \end{bmatrix}$$

#### Character table for the dihedral group $D_6$

Now recall that  $D_6$  is basically the nontrivial semidirect product of  $\mathbb{Z}_3$  and  $\mathbb{Z}_4$ .

Now one dimensional representation are given by

$$(A, B, C) \rightarrow \{\pm 1\}, \text{ where } D_6 = \langle A, B, C \rangle.$$

Now  $A$  acts by a diagonal scalar matrix, thus two dimensional representations  $\rho_0, \dots, \rho_3$  split into one-dimensional representations. On the other hand the representations  $\rho_0, \rho_1$  and  $\rho_2$  are irreducible.

Hence the character table for  $D_6$  is

$$\begin{bmatrix} & 1 & 2A & 2B & 2C & 3 & 6 \\ \rho_0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \rho_1 & 1 & 1 & -1 & -1 & 1 & 1 \\ \rho_2 & 1 & -1 & -1 & 1 & 1 & -1 \\ \rho_3 & 1 & -1 & 1 & -1 & 1 & -1 \\ \rho_4 & 2 & 2 & 0 & 0 & -1 & -1 \\ \rho_5 & 2 & -2 & 0 & 0 & -1 & 1 \end{bmatrix}$$



### Character table for the non-abelian group of order 12

Let  $G$  be the non-abelian group of order 12. Notice that

$$G = \{a, b \mid a^7 = b^3 = e, bab^{-1} = a^2\}.$$

Note that there are five classes of  $G$ . They are given by the class equation

$$|G| = 1 + 1 + 3 + 3 + 7 + 7.$$

Now  $\langle a \rangle$  is normal in  $G$ . Therefore the irreducible representations of  $G/\langle a \rangle \approx \mathbb{Z}_3$  gives 3 irreducible representation of degree one.

Let us assume  $\rho_4$  and  $\rho_5$  are remaining irreducible representations of degree  $e$  and  $f$ . Then we have

$$21 = 3 + e^2 + f^2.$$

By using the orthogonal properties, the character table for  $G$  is given by

$$\begin{bmatrix} & e & a & a^3 & b & b^3 \\ \chi_1 & 1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & 1 & 1 & w & w^2 \\ \chi_3 & 1 & 1 & 1 & w^2 & w \\ \chi_4 & 3 & \bar{z} & z & 0 & 0 \\ \chi_5 & 3 & z & \bar{z} & 0 & 0 \end{bmatrix}$$

where  $w = \frac{-1+i\sqrt{3}}{2}$  and  $z = \frac{-1+i\sqrt{7}}{2}$ .

### Result

5

We have find the character tables for the group  $V, Q, D_4, D_6$  and  $G$ , non-abelian group of order 12.

### 4. a

We have  $G = D_5$  given by the presentation  $x^5 = 1 = y^2, yx = x^{-1}y, D_5 = \{1, x, x^2, x^3, x^4, y, yx, yx^2, yx^3, yx^4\}$ . Now, the conjugacy class is well known and is as follows

$$\{1\}, \{x, x^4\}, \{x^2, x^3\}, \{x, xy, xy^2, xy^3, xy^4\}$$

We see that there are exactly 4 conjugacy classes, and therefore we get, there are four distinct irreducible characters of  $D_5$ . Let  $\chi_i$  denote the characters of  $D_5, 1 \leq i \leq 4$ .

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 = 10$$

, where  $d_i (1 \leq i \leq 4)$  are the degrees of  $\chi_i$ .

Observe that  $d_i \leq 4$ , for every  $1 \leq i \leq 4$ . Observe that none of them can't be 3, because if they do, then,

$$d_1^2 + d_2^2 + d_3^2 = 1$$

which is not possible, as each of them are positive integers. We conclude that the degrees are only 1 and 2. All of them can't be one, and all of them can't be 2. Therefore, there are 3 choices remaining, one of them is that 3 representations have degree 1, and the other of 2, but squares of this combination add upto 6, which is contradiction. Also, three of them can't be 2 for the same reason. The last remaining choice is two of them are degree 1, and the other two are of 2, and this combination indeed matches the equation. Therefore, we conclude,

$$d_1 = d_2 = 1, d_3 = d_4 = 2$$

Now, we have information about the possible degrees of irreducible characters of  $D_5$ .

Now, we proceed to solve the problem. Let  $\chi$  be a character of degree 2. Now,  $x^5 = 1$ . Suppose  $\rho$  is the representation that affords  $\chi$ . Since  $\rho$  is a homomorphism, we have  $\rho(x)^5 = 1$ . This implies that the minimal polynomial of  $\rho(x)$  divides  $x^5 - 1$ , and hence the eigen values of  $\rho(x)$  are fifth roots of unity. Also, since  $x^5 - 1$  has 5 distinct roots, this implies that  $\rho(x)$  has two distinct roots which are fifth root of unity. Since  $\chi(x) = \text{trace}(\rho(x))$ , which in turn is the sum of eigen values we conclude that  $\chi(x)$  is the sum of two distinct fifth roots of unity. This answers (a).

We have that  $x$  and  $x^{-1} = x^4$  are conjugate. Since  $\chi$  is a class function we have  $\chi(x) = \chi(x^{-1})$ . But we know that  $\chi(x^{-1}) = \overline{\chi(x)}$ . This is proposition 10.4.8 (d). So, we have that  $\chi(x) = \overline{\chi(x)}$ . Hence we conclude that  $\chi(x) \in \mathbb{R}$ . This solves (b).

Next, we have to determine the character table of  $D_5$ . To do this, first we determine the linear characters. There are only two linear characters, which has been proved before. Let  $\chi_1$  be the trivial character. Now, let  $\chi_2$  be another degree one character. Observe that  $\chi_2(x)$  is a fifth root of unity. The reason being exactly same as proved in (a). Also since  $\chi_2(x) = \chi_2(x^{-1})$ , we have  $\chi_2(x) = 1$ . The exact same reasoning holds for  $\chi_2(x^2) = \chi_2(x^3)$ . Therefore  $\chi_2(x^2) = 1$ . Now,  $\chi_2(y)^2 = 1 \implies \chi_2(y) = \pm 1$ . But since  $\chi_2$  is not the trivial character we conclude that  $\chi_2(y) = -1$ . This determines the degree one characters.

To determine the degree 2, characters we will again use (a) and (b).  $\chi_3, \chi_4$  be the characters of degree 2. Now observe that  $\chi_3(y)$  is the sum of two distinct squares of unity. This follows from the proof of (a) along with the fact that  $y^2 = 1$ . But  $\pm 1$  are the only squares of unity, and therefore  $\chi_3(y) = 1 - 1 = 0$ . Now, from (b), we have  $\chi_3(x) \in \mathbb{R}$ , and also from (a), we have that  $\chi(x)$  is the sum of two distinct fifth roots of unity. Since,  $1, e^{i\pi/5}, e^{2i\pi/5}, e^{3i\pi/5}, e^{4i\pi/5}$  are the only fifth roots of unity, we have two possibilities of  $\chi_3(x)$ .  $\chi_3(x) = e^{\pi/5} + e^{4\pi/5} = 2 \cos \pi/5$  or,  $\chi_3(x^2) = e^{2\pi/5} + e^{3\pi/5} = 2 \cos 2\pi/5$ . Let  $\chi_3(x) = 2 \cos \pi/5$ . Now, we have to determine  $\chi_3(x^2)$ . Now, observe that  $\langle \chi_1, \chi_3 \rangle = 1$ . This gives us the follows

$$\langle \chi_1, \chi_3 \rangle = 1 \implies \frac{1}{10}(1.2 + 2.2 \cos \pi/5 + 2. \chi_3(x^2) + 5.0) = 1$$

$$\chi_3(x^2) = 4 - 2 \cos \pi/5 = 2 \cos 2\pi/5$$

This completely determines  $\chi_3$ . Exactly similar reasoning for  $\chi_4$  will say that  $\chi_4(x) = 2 \cos(2\pi/5), \chi_4(x^2) = 2 \cos \pi/5, \chi_4(y) = 0$ .

The following is the character table of  $D_5$

	1	$x$	$x^2$	$y$
$\chi_1$	1	1	1	1
$\chi_2$	1	1	1	-1
$\chi_3$	2	$2 \cos \pi/5$	$2 \cos 2\pi/5$	0
$\chi_4$	2	$2 \cos 2\pi/5$	$2 \cos \pi/5$	0



Now, we finish the problem by doing (d), which is easy.  $C_5 = \{1, x, x^2, x^3, x^4\}$ . Now  $\chi_1$  and  $\chi_2$  is easily seen to restrict to the trivial character of  $C_5$

$$\text{Now } \chi_3(1) = 2, \chi_3(x) = \chi_3(x^4) = 2 \cos \pi/5, \chi_3(x^2) = \chi_3(x^3) = 2 \cos 2\pi/5$$

$$\text{Define } \rho_1 : C_5 \rightarrow \mathbb{C}^* \text{ defined by } \rho_1(x) = e^{\pi/5}. \text{ Define } \rho_2 : C_5 \rightarrow \mathbb{C}^* \text{ defined by } \rho_1(x) = e^{-\pi/5}.$$

Then it is clear that  $\chi_3 = \rho_1 + \rho_2$ .

Lastly,

$$\text{Now } \chi_4(1) = 2, \chi_4(x) = \chi_4(x^4) = 2 \cos 2\pi/5, \chi_4(x^2) = \chi_4(x^3) = 2 \cos(-2\pi/5)$$

$$\text{Define } \rho_3 : C_5 \rightarrow \mathbb{C}^* \text{ defined by } \rho_3(x) = e^{2\pi/5}. \text{ Define } \rho_4 : C_5 \rightarrow \mathbb{C}^* \text{ defined by } \rho_4(x) = e^{-2\pi/5}.$$

Then it is clear that  $\chi_4 = \rho_3 + \rho_4$ .

This completely solves the problem.

## Result

The following is the character table of  $D_5$

	1	$x$	$x^2$	$y$
$\chi_1$	1	1	1	1
$\chi_2$	1	1	1	-1
$\chi_3$	2	$2 \cos \pi/5$	$2 \cos 2\pi/5$	0
$\chi_4$	2	$2 \cos 2\pi/5$	$2 \cos \pi/5$	0

See the solution for more details

## 5. a

Consider the group following group:

$$G = \langle x, y \mid x^5 = y^4 = e, yxy^{-1}x^{-2} \rangle$$

Find the character table of  $G$ .

[Comment](#)

Step 2 of 5 ^

The group  $G$  of order 20 with  $G = \langle x, y \mid x^5 = y^4 = e, yxy^{-1}x^{-2} \rangle$

It has the conjugacy classes as shown below,

$$\{e\}, \{x, x^2, x^3, x^4\}, \{yx^m\}, \{y^2x^m\}, \{y^3x^m\}$$

Since  $yxy^{-1} = x^{-2}$  this implies that  $yxy^{-1}x^{-1} = x^{-1}$ .

So the commutator subgroup contains  $x$  that is, it is equal to  $\langle x \rangle$ .

It is known that the subgroup generated by  $\langle x \rangle$  is normal and a quotient group  $G/\langle x \rangle$  is isomorphic to  $Z/4$ .

Therefore, there are 5 irreducible representations, 4 of which are 1-dimensional.

Suppose that the last one is  $d$ -dimensional,

Then,

$$20 = 1^2 + 1^2 + 1^2 + 1^2 + d^2$$

This implies that,

$$20 = 1 + 1 + 1 + 1 + d^2$$

$$20 = 4 + d^2$$

$$20 - 4 = d^2$$

$$16 = d^2$$

This implies that,

$$d = 4$$

The representation of 4 one dimensional as shown below,

$$x^m y^k \rightarrow e^{2\pi i km/4} \text{ For } m = 0, 1, 2, 3$$

This gives the first 4 rows of the character table as shown below,

	$e$	$x^m$	$yx^m$	$y^2x^m$	$y^3x^m$
$\chi_1$	1	1	1	1	1
$\chi_2$	1	1	$i$	$-1$	$-i$
$\chi_3$	1	1	$-1$	1	$-1$
$\chi_4$	1	1	$-i$	$-1$	$i$
$\chi_5$	4	$a$	$b$	$c$	$d$

Now find the last row by using the character orthogonally relations.

$$\langle \chi_1 | \chi_5 \rangle = 4 + a + b + c + d = 0 \quad \dots (1)$$

$$\langle \chi_2 | \chi_5 \rangle = 4 + a + ib - c - id = 0 \quad \dots (2) \quad \langle \chi_3 | \chi_5 \rangle = 4 + a - b + c - d = 0 \quad \dots (3)$$

$$\langle \chi_4 | \chi_5 \rangle = 4 + a - ib - c + id = 0 \quad \dots (4)$$

Find the  $a, b, c, d$  by using four equation.

Equation (1) and (3), to get

$$8 + 2a + 2c = 0 \quad \dots (5)$$

Equation (2) and (4), to get

$$8 + 2a - 2c = 0 \quad \dots (6)$$

Now, solve equation (5) and equation (6), to get

$$16 + 4a = 0$$

$$4a = -16$$

$$a = -4$$

Substitute  $a = -4$  into equation (6), to get

$$8 + 2(-4) - 2c = 0$$

$$8 - 8 - 2c = 0$$

$$c = 0$$

Substitute  $a = -4, c = 0$  into equation (1) and equation (2),

$$b + d = 0$$

And,

$$ib - id = 0$$

Solve the above two equation, to get  $b$  and  $d$

$$b = 0$$

$$d = 0$$

Therefore, the character table of  $G$  as shown below,

	$e$	$x^m$	$yx^m$	$y^2x^m$	$y^3x^k$
$\chi_1$	1	1	1	1	1
$\chi_2$	1	1	$i$	$-1$	$-i$
$\chi_3$	1	1	$-1$	1	$-1$
$\chi_4$	1	1	$-i$	$-1$	$i$
$\chi_5$	4	$-4$	0	0	0

The group  $G$  of order 20  $\langle x, y \mid x^5 = y^4 = e, yxy^{-1} = x^2 \rangle$  has the following conjugacy classes.

We also have  $[y, x] = yxy^{-1}x^{-1} = x$ , so the commutator subgroup contains  $x$ ; in fact, it equals  $\langle x \rangle$ . The subgroup generated by  $x$  is normal and the quotient  $G/\langle x \rangle$  is isomorphic to  $Z/4$ .

Therefore, there are 5 irreducible representations, 4 of which are 1-dimensional. If the last one is  $d$ -dimensional, we must have  $20 = 1^2 + 1^2 + 1^2 + 1^2 + d^2$ , so  $d = 4$ . The four 1-dimensional representations are given by  $x \mapsto \zeta^k, y \mapsto 1$  for  $k = 0, 1, 2, 3$ . This gives us the first 4 rows of the character table. The last row we can cheat and get for free by using the character orthogonality relations.  $e \ x \ y \ x^2 \ y^2 \ x^3 \ y^3 \ x^4 \ y^4 \ x^5 \ y^5 \ x^6 \ y^6 \ x^7 \ y^7 \ x^8 \ y^8 \ x^9 \ y^9 \ x^{10} \ y^{10}$

## 6. a

First we show that the columns of character table are orthogonal. Let  $G$  be the group, and let  $C_1, \dots, C_k$  be the conjugacy classes of  $G$ . Therefore, there are  $k$  distinct characters  $\chi_1, \dots, \chi_k$ . Observe that if  $A = (a_{ij})$  is the character table, then the  $i^{th}$  column  $(a_{ij})_{1 \leq j \leq k}$ , is given by  $(\chi_i(C_j))_{1 \leq j \leq k}$ . Note that since each character is a class function, there is no harm in writing the notation  $\chi(C)$  for some conjugacy class  $C$ , and some character  $\chi$ .

Now, Let  $C, C'$  be two conjugacy classes. For a conjugacy class  $C$ , let  $\delta_C$ , denote the class function on  $G$ , defined as follows

$$\delta_C(g) = \begin{cases} 1 & , g \in C \\ 0 & , otherwise \end{cases}$$

Now, since the distinct irreducible characters  $\chi_i (1 \leq i \leq k)$ , form a set of orthogonal basis for the space of class function on  $G$ , we can write

$$\delta_C = \sum_{i=1}^k \langle \delta_C, \chi_i \rangle \chi_i = \frac{|C|}{|G|} \sum_{i=1}^k \chi_i(C) \chi_i.$$

$$\text{Similarly, } \delta_{C'} = \sum_{i=1}^k \langle \delta_{C'}, \chi_i \rangle \chi_i = \frac{|C'|}{|G|} \sum_{i=1}^k \chi_i(C') \chi_i.$$

Now, observe that,

$$\langle \delta_C, \delta_{C'} \rangle = \begin{cases} 0 & \text{if } C \neq C' \\ \frac{|C|}{|G|} & \text{if } C = C' \end{cases}$$

Now as,

$$\langle \chi_i, \chi_j \rangle = \begin{cases} 0 & , \text{if } i \neq j \\ 1 & , \text{if } i = j \end{cases}$$

we have that  $\langle \delta_C, \delta_{C'} \rangle = \frac{|C||C'|}{|G|^2} \sum_{i=1}^k \overline{\chi_i(C)} \chi_i(C')$

Hence if  $a_j$ , and  $a_l$  are two columns of the character table  $A$ , then

$\langle a_j, a_k \rangle = \sum_{i=1}^k \chi_i(C_j) \overline{\chi_i(C_l)}$ , and we have

$$\langle a_j, a_k \rangle = \begin{cases} 0 & \text{if } j \neq l \\ \frac{|G|}{|C_j|} & \text{if } j = l \end{cases}$$

This proves that the columns of the character table are orthogonal

This also proves that

$$AA^* = \begin{bmatrix} \frac{|G|}{|C_1|} & & & \\ & \frac{|G|}{|C_2|} & & \\ & & \ddots & \\ & & & \frac{|G|}{|C_k|} \end{bmatrix}$$

So, we can make the following adjustment to the matrix  $A$ . To each column  $a_i$  of  $A$ , we can multiply the term  $\frac{\sqrt{|C_i|}}{\sqrt{|G|}}$ . In that way we observe that  $AA^* = I$ , and hence  $A$  becomes unitary.

## Result

3 of 3

We first prove that the columns of the character table  $A$  of  $G$ , are orthogonal. From this fact, we mention the necessary adjustment the matrix  $A$ , need to have, so that it becomes a unitary matrix

## 7. a

Suppose  $G$  is a finite group and  $N$ , a normal subgroup of  $G$ . Let  $\rho' : G/N \rightarrow GL(V)$ , be a irreducible representation of  $G/N$ . Let  $\pi : G \rightarrow G/N$ , denote the canonical map. Consider now, the map  $\rho = \rho' \circ \pi : G \rightarrow GL(V)$ . So,  $\rho(g) = \rho'(gN)$ , for every  $g \in G$ . We prove that  $\rho$ , is an irreducible representation of  $G$ . Suppose  $W \neq 0$ , be a  $G$ -invariant subspace of  $V$ . So, for every  $g \in G$ , and  $w \in W$ , we have  $\rho(g)(w) \in W$ . So,  $\rho'(gN)(w) \in W$ . We conclude that  $W$  is  $G/N$  invariant subspace of  $V$ . But, as  $\rho'$  is irreducible  $W = V$ . So,  $\rho$  is irreducible.

Now, we prove the same result using theorem 10.4.6. Let  $\chi'$  be the character of  $\rho'$ , and  $\chi$  the character of  $\rho$ . Observe that for  $g \in G$ ,  $\chi(g) = \chi'(gN)$ . So, if  $g_1N = g_2N$ , we have  $\chi(g_1) = \chi(g_2)$ . Now, since  $\rho'$  is irreducible, we have  $\langle \chi', \chi' \rangle = 1$ . Expanding that we have,

Now, we prove that  $\langle \chi, \chi \rangle = 1$ . Indeed, we have  $\langle \chi, \chi \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\chi(g)}$ .

$$\frac{|N|}{|G|} \sum_{g \in G/N} \chi'(g) \overline{\chi'(g)} = 1 \quad (1)$$

But, we observe that if  $g \in aN$ , then  $\chi(g) = \chi'(aN)$ .

$$\frac{|N|}{|G|} \sum_{g \in G/N} \chi'(g) \overline{\chi'(g)} = 1 \quad (2)$$

So,  $\sum_{g \in aN} \chi(g) \overline{\chi(g)} = |N| \chi'(aN) \overline{\chi'(aN)}$ . Hence, we have  $\sum_{g \in G} \chi(g) \overline{\chi(g)} = \sum_{g \in G/N} |N| \chi'(g) \overline{\chi'(g)} = \frac{|G|}{|N|}$ . The last equality follows from equation (1). So,  $\langle \chi, \chi \rangle = \frac{|N|}{|G|} \cdot \frac{|G|}{|N|} = 1$ . Therefore,  $\rho$  is irreducible.

## Result

3 of 3

The idea is to compute the character  $\chi'$  of the representation  $\rho'$  of  $G/N$ . We actually determine  $\langle \chi, \chi \rangle$ . See the solution for more details.

8. a

**Solution:** Given below the incomplete character table as

$$\begin{array}{c|ccccc} & (1) & (3) & (6) & (6) & (8) \\ \hline \chi_1 & 1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & 1 & -1 & -1 & 1 \\ \chi_3 & 3 & -1 & 1 & -1 & 0 \\ \chi_4 & 3 & -1 & -1 & 1 & 0 \end{array}$$

Let the missing character be  $\chi_5$  and the values are  $a, b, c, d, e$ .

Therefore we have to fill up the below table as

$$\begin{array}{c|ccccc} & (1) & (3) & (6) & (6) & (8) \\ \hline \chi_1 & 1 & 1 & 1 & 1 & 1 \\ \chi_2 & 1 & 1 & -1 & -1 & 1 \\ \chi_3 & 3 & -1 & 1 & -1 & 0 \\ \chi_4 & 3 & -1 & -1 & 1 & 0 \\ \chi_5 & a & b & c & d & e \end{array}$$

Now notice that the group has five conjugacy classes. It has five different irreducible representations, so exactly one row is missing.

Now notice that the group has 24 elements since

$$1 + 3 + 6 + 6 + 8 = 24.$$

Therefore we have

$$\begin{aligned} 1^2 + 1^2 + 3^2 + 3^2 + a^2 &= 24 \\ \implies a &= 2. \end{aligned}$$

This follows that

$$\dim V_5 = 2.$$

Similarly by orthogonality of characters we have

$$a + 3b + 6c + 6d + 8e = a + 3b - 6c - 6d + 8e = 0.$$

This follows that

$$a + 3b + 8e = 6c + 6d = 0.$$

Hence it follows that

$$3a - 3b + 6c - 6d = 3a - 3b - 6c + 6d = 0.$$

Therefore note that

$$3a - 3b = 6c - 6d = 0.$$

Thus we have

$$c = d = 0, \quad \text{since } c + d = c - d = 0.$$

Also we have

$$a = b = 2.$$

Now for  $c$  look that

$$c = -\frac{(a + 3b)}{8} = -1.$$

Therefore the required solution is

$$(a, b, c, d, e) = (2, 2, -1, 0, 0).$$

This completes the solution.

---

## Result

The missing values  $a, b, c, d, e$  are  $2, 2, -1, 0, 0$  respectively.

9. a



Below is a partial character table. One conjugacy class is missing. Hence, one irreducible character is also missing. So, after inserting the last column and last row, and replacing them by variables, we get the following,

	(1)	(1)	(2)	(2)	(3)	
	1	$u$	$v$	$w$	$x$	$z$
$\chi_1$	1	1	1	1	1	1
$\chi_2$	1	1	1	1	-1	$a_2$
$\chi_3$	1	-1	1	-1	$\iota$	$a_3$
$\chi_4$	1	-1	1	-1	$-\iota$	$a_4$
$\chi_5$	2	2	-1	-1	0	$a_5$
$\chi_6$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$a_6$

We have that  $\chi_1$  is the trivial character and hence the last entry of the first row is 1. The unknown conjugacy class has a representative  $z$ (say). We have to determine all the variables of the character table, along with the order of the conjugacy class of  $z$ .

We prove the following simple result which will be of use. Let  $\chi : G \rightarrow \mathbb{C}^*$  be a one dimensional representation of  $G$ , and hence a character of degree 1. Let  $R : G \rightarrow GL(V)$  be any other representation of  $G$ . Let  $\chi_R$  be its character. Define the following representation  $\rho : G \rightarrow GL(V)$  as

$$\rho_g = \chi(g)R_g$$

It is clear that  $\rho$  is a homomorphism(routine check). Let  $\chi_\rho$  be its character. Then,

$$\chi_\rho(g) = \chi(g)\chi_R(g)$$

Now, we claim the following,

**Claim:** If  $R$  is irreducible then so is  $\rho$ .

$$\text{Now, } \langle \chi_\rho, \chi_\rho \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g) \overline{\chi_\rho(g)} = \frac{1}{|G|} \sum_{g \in G} \chi(g) \chi_R(g) \overline{\chi(g) \chi_R(g)} = \frac{1}{|G|} \sum_{g \in G} |\chi(g)|^2 \chi_R(g) \overline{\chi_R(g)} = \frac{1}{|G|} \sum_{g \in G} \chi_R(g) \overline{\chi_R(g)} = \langle \chi_R, \chi_R \rangle = 1.$$

Observe that  $|\chi(g)| = 1$ , as  $\chi(g)$  is a root of unity in  $\mathbb{C}^*$ . Then last equality follows as  $R$  is irreducible.

Now, we will apply this to our situation to find the remaining character. Consider the one dimensional character  $\chi_4$ , and the two dimensional character  $\chi_5$ . Then by the above the result that has been proved, we have that  $\chi = \chi_4 \chi_5$  is an irreducible character of degree 2. Now, to say that  $\chi$  is actually  $\chi_6$ , we need to see whether  $\chi$  is distinct from  $\chi_5$ . But that is true as  $\chi(u) = \chi_4(u) \chi_5(u) = -2 \neq \chi_5(u)$ . So, we have now identified the remaining character  $\chi_6 = \chi_4 \chi_5$ . Therefore,

$$\chi_6(1) = 2, b_2 = \chi_6(u) = -2, b_3 = \chi_6(v) = -1, b_4 = \chi_6(w) = 1, b_5 = \chi_6(x) = 0$$

Observe that  $|G| = \sum_{i=1}^6 \chi_i(1)^2 = 12$ . So, we have that  $|Cl(z)| = 3$ , where  $Cl(z)$  denote the conjugacy class of  $z$ . It remains to determine  $a_i$  ( $2 \leq i \leq 6$ ). To determine that observe that  $\langle \chi_1, \chi_i \rangle = 1$  for every  $2 \leq i \leq 6$ . We find each indeterminate one by one.

$$\begin{aligned} \langle \chi_1, \chi_2 \rangle = 1 &\implies \frac{1}{12}(1.1 + 1.1 + 2.1 + 2.1 + 3. - 1 + 3.a_2) = 1 \\ &\implies 1 + a_2 = 4 \implies a_2 = 3 \end{aligned}$$

$$\begin{aligned} \langle \chi_1, \chi_3 \rangle = 1 &\implies \frac{1}{12}(1.1 + 1. - 1 + 2.1 + 2. - 1 + 3.\iota + 3.a_3) = 1 \\ &\implies \iota + a_3 = 4 \implies a_3 = 4 - \iota \end{aligned}$$

$$\begin{aligned} \langle \chi_1, \chi_4 \rangle = 1 &\implies \frac{1}{12}(1.1 + 1. - 1 + 2.1 + 2. - 1 + 3. - \iota + 3.a_4) = 1 \\ &\implies -\iota + a_4 = 4 \implies a_4 = 4 + \iota \end{aligned}$$

$$\begin{aligned} \langle \chi_1, \chi_5 \rangle = 1 &\implies \frac{1}{12}(1.2 + 1.2 + 2. - 1 + 2. - 1 + 3.0 + 3.a_5) = 1 \\ &\implies a_5 = 4 \end{aligned}$$

$$\begin{aligned} \langle \chi_1, \chi_6 \rangle = 1 &\implies \frac{1}{12}(1.2 + 1. - 2 + 2. - 1 + 2.1 + 3.0 + 3.a_6) = 1 \\ &\implies a_6 = 4 \end{aligned}$$

We have now determined all the variable in the character table above. Therefore, we write the completed table as follows:

	(1)	(1)	(2)	(2)	(3)	(3)
	1	$u$	$v$	$w$	$x$	$z$
$\chi_1$	1	1	1	1	1	1
$\chi_2$	1	1	1	1	-1	3
$\chi_3$	1	-1	1	-1	$\iota$	$4 - \iota$
$\chi_4$	1	-1	1	-1	$-\iota$	$4 + \iota$
$\chi_5$	2	2	-1	-1	0	4
$\chi_6$	2	-2	-1	1	0	4

Now, we determine the order of the representatives of each conjugacy class. Observe that  $|Cl(u)| = 1$ . The sizes of all the other conjugacy classes are all more than 1. Therefore we conclude that  $Z(G) = \{1, u\}$ . This also says that  $o(u) = 2$ . Next consider  $\text{Ker}\chi_3 = 1 \cup Cl(v)$ . This is clearly a normal subgroup of  $G$ . But,  $|Cl(v)| = 2$ , implies that  $\text{Ker}\chi_3$  has order 3, hence cyclic. This implies  $o(v) = 3$ . We have  $|Cl(x)| = 3 \implies |C(x)| = 4$ , where  $C(x)$  denote the centralizer of  $x$ . Now, observe that  $u \in C(x)$ , as  $u$  is in the center of  $G$ . So,  $ux \in C(x)$ , and  $ux \neq u \neq x$ . This implies  $C(x) = \{1, u, x, ux\}$ . Therefore,  $C(x) \cong C_2 \times C_2$ , and hence  $o(x) = 2$ . Using the exact same argument, we conclude that  $o(z) = 2$ . The only element that remain is  $w$ . Now, observe that  $o(u) = 2, o(v) = 3$ , and  $uw = vw$ , implies that  $o(uv) = o(u)o(v) = 6$ . This is a standard result in group theory. Now, since all the other conjugacy classes doesn't have element of order 6, we conclude that  $uv \in Cl(z)$ , and hence  $o(w) = 6$ . So, to conclude  $o(u) = o(x) = o(z) = 2, o(w) = 6, o(v) = 6$ .

Now, we determine all normal subgroup of  $G$ . To determine we use the following general principle that we have been using in other problems also. Given a finite group  $G$ , and its character table  $T$ , how can one find all normal subgroups of  $G$ . To do this we first prove the following claim:

**Claim 1** Suppose  $R$  is a representation of  $R : G \rightarrow GL(V)$  of degree  $d$ , and  $\chi_R$  denote the character of  $R$ . Define  $Ker(\chi_R) = \{g \in G | \chi_R(g) = d\}$ . We claim that  $Ker(\chi_R) = Ker R$ , where  $Ker R$  is the usual kernel of the homomorphism  $R$ . Indeed if  $g \in Ker R$ , then  $R_g = id_V$ . Since degree of  $R$  is  $d$ , we get  $\chi_R(g) = \text{trace}(id_V) = d$ . Therefore  $g \in Ker(\chi_R)$ . This proves  $Ker R \subseteq Ker(\chi_R)$ .

Conversely suppose  $g \in Ker(\chi_R)$ . This implies  $\chi_R(g) = d \implies |\chi_R(g)| = d$ . Now, it is a standard observation that  $\chi_R(g)$  is sum of  $d$  roots of unity. Also, it is a standard result that if  $\zeta_1, \dots, \zeta_d$  are roots of unity then  $|\zeta_1 + \dots + \zeta_d| \leq d$ , with equality iff  $\zeta_1 = \zeta_2 = \dots = \zeta_d$ . so, from this we conclude that as  $|\chi_R(g)| = d$ , then  $\chi_R = d\zeta$ , where  $\zeta$  is a root of unity. But,  $\chi_R(g) = d$  implies that  $\zeta = 1$ . Hence, we get that all the eigen values of  $R_g$  is 1. But  $R_g$  is diagonalizable (again standard observation, as minimal polynomial of  $R_g$  divides  $X^n - 1$ , where  $n$  is the order of  $g$ ), and therefore we conclude that  $R_g = id_V$ . Therefore,  $Ker(\chi_R) \subseteq Ker R$ . This completes the proof of the claim.

Suppose now that  $\chi$  is any character (not necessarily irreducible) of  $G$ . Let  $\chi = \sum_{i=1}^k n_i \chi_i$ , where  $\chi_1, \chi_2, \dots, \chi_k$  denote all the distinct irreducible characters, and  $n_i (1 \leq i \leq k)$  are non-negative integers. Let  $d_i$  denote the degrees of the character  $\chi_i$  for every  $1 \leq i \leq k$ .

**Claim 2:**  $Ker \chi = \cap \{ Ker \chi_i | n_i > 0 \}$ .

Let  $A = \cap \{ Ker \chi_i | n_i > 0 \}$ . Let  $g \in A$ . Then  $\chi_i(g) = d_i$ , for every  $i$  such that  $n_i > 0$ . If  $d$  is the degree of  $\chi$ , it is clear that  $\chi(g) = d$ . Therefore,  $g \in Ker \chi$ . This proves  $A \subseteq Ker \chi$ . Conversely, suppose  $g \in Ker \chi$ , then  $\chi(g) = d$ . So  $|\chi_i(g)| \leq d_i$ , claim 1 forces  $\chi_i(g) = d_i$ , whenever  $n_i > 0$ . Therefore, we have that  $g \in A$ . Thus, we have proved our second claim.

Now, we proceed towards the final claim, which tells us how to determine all normal subgroups of  $G$ , by the character table of  $G$ . Let,  $N_i = Ker \chi_i$ . It is clear the  $N_i$  are normal subgroups.

**Claim 3:** Any normal subgroup of  $G$  is a certain intersection of  $N_i$ 's.

Let  $N$  be any normal subgroup of  $G$ . Consider  $G/N$ . Let  $\rho$  denote the regular representation of  $G/N$ . Let  $\pi : G \rightarrow G/N$  be the canonical surjection, with  $Ker \pi = N$ . Consider  $\rho_G = \rho \circ \pi$ . Observe that since  $\rho$  is injective,  $\rho_G$  is a representation of  $G$ , with  $Ker \rho_G = N$ . Let  $\chi_G$  be the character of  $\rho_G$ . Then, by claim 1,  $N = Ker \chi_G$ . Now, Let

$$\chi_G = \sum_{i=1}^k n_i \chi_i$$

. Then, from claim 2, we get  $N = Ker \chi_G = \cap \{ Ker \chi_i | n_i > 0 \} = \cap \{ N_i | n_i > 0 \}$ .

So, finally it is easy to determine the  $N_i$ 's from the character table. To determine  $N_i$ , just look at the  $i^{th}$  row, and find out all elements of  $G$ , whose character entry is  $d_i$ , where  $d_i = \chi_i(1)$ . Then to determine all normal subgroups, just take all possible intersection of these  $N_i$ 's, and list the new subgroups thus obtained. This gives you all the possible normal subgroups.

Now, we determine all normal subgroup of  $G$ . To determine we use the following general principle that we have been using in other problems also. Hence, using this,

$$\begin{aligned} N_1 &= \text{Ker}\chi_1 = G \\ N_2 &= \text{Ker}\chi_2 = 1 \cup Cl(u) \cup Cl(v) \cup Cl(w) \\ N_3 &= \text{Ker}\chi_3 = \{1, v, v^2\} \\ N_4 &= \text{Ker}\chi_4 = \{1, v, v^2\} \\ N_5 &= \text{Ker}\chi_5 = \{1, u\} \\ N_6 &= \text{Ker}\chi_6 = \{1\} \end{aligned}$$

So, taking all possible intersection of these  $N_i$ 's doesn't yield any new subgroup. Therefore the distinct normal subgroups are  $1, G, Z(G) = \{1, u\}, \{1, v, v^2\}$  and  $\{1, u\} \cup Cl(v) \cup Cl(w)$ .

(d) Next we have to determine the structure of the group. We observe that  $|G| = 12 = 2^2 \cdot 3$ . Observe that there is a normal subgroup of order 3 say  $H = \text{Ker}\chi_3$ . Now, Consider a 2-sylow subgroup  $K$  of order 5. Observe that  $G = HK$ , and  $H \cap K = 1$ . Hence  $G \cong H \rtimes K$ .

## Result

The following is the complete character table. See the answer for the remaining parts and the relevant explanations.

	(1)	(1)	(2)	(2)	(3)	(3)
	1	$u$	$v$	$w$	$x$	$z$
$\chi_1$	1	1	1	1	1	1
$\chi_2$	1	1	1	1	-1	3
$\chi_3$	1	-1	1	-1	$\iota$	$4 - \iota$
$\chi_4$	1	-1	1	-1	$-\iota$	$4 + \iota$
$\chi_5$	2	2	-1	-1	0	4
$\chi_6$	2	-2	-1	1	0	4

10. a

We have a character table of a group  $G$ , with missing rows. There are five columns, which means there are five conjugacy classes and hence forth there are five distinct irreducible characters. There are four rows representing four distinct characters  $\chi_1, \chi_2, \chi_3, \chi_4$ . Therefore there is only one remaining row which needs to be determined. For, that we name the fifth character  $\chi_5$ . We need to determine this character on representative of the conjugacy classes already written in the table.

First we determine the degree of  $\chi_5$ . Let  $d_i$  denote the degree  $\chi_i$ , for  $1 \leq i \leq 5$ . Observe that  $d_1 = d_2 = d_3 = d_4 = 1$ . Also,  $\sum_{i=1}^5 d_i^2 = 20 \implies d_5 = 4$ . Now, Let  $\chi_G$  denote the character of the regular representation of  $G$ . Then

$$\chi_G(g) = \begin{cases} 20 & \text{if } g = e \\ 0 & , \text{ otherwise} \end{cases}$$

Now we also know  $\chi_G = \sum_{i=1}^5 d_i \chi_i = \chi_1 + \chi_2 + \chi_3 + \chi_4 + 4\chi_5 = \chi_5 \implies \frac{1}{4} [\chi_G - \chi_1 - \chi_2 - \chi_3 - \chi_4]$

$$\chi_5(a) = \frac{1}{4} [0 - 1 - 1 - 1 - 1] = -1$$

$$\chi_5(b) = \frac{1}{4} [0 - 1 + 1 + i - i] = 0$$

$$\chi_5(c) = \frac{1}{4} [0 + 1 - 1 + i - i] = 0$$

$$\chi_5(d) = \frac{1}{4} [0 - 1 - 1 + 1 + 1] = 0$$

Hence the final character table is as follows:

	(1)	(4)	(5)	(5)	(5)
	0	a	b	c	d
$\chi_1$	1	1	1	1	1
$\chi_2$	1	1	-1	-1	1
$\chi_3$	1	1	$-\iota$	$\iota$	-1
$\chi_4$	1	1	$\iota$	$-\iota$	-1
$\chi_5$	4	-1	0	0	0

Next, we determine the order of  $a, b, c, d$ . Let  $C(a)$  denote the centralizer of  $a$  in  $G$ . Let  $Cl(a)$  denote the conjugacy class of  $a$  in  $G$ . We know,  $|Cl(a)| = \frac{|G|}{|C(a)|} \implies |C(a)| = 5$ . Now  $a \in C(a)$ , and since order of  $C(a)$  is prime,  $C(a)$  is cyclic and hence  $|a| = 5$ .

Let  $C(b)$  denote the centralizer of  $b$  in  $G$ . Let  $Cl(b)$  denote the conjugacy class of  $b$  in  $G$ . We know,  $|Cl(b)| = \frac{|G|}{|C(b)|} \implies |C(b)| = 4$ . Since  $b \in C(b)$ , it is clear  $|b| = 2$  or  $4$ . But, if  $|b| = 2$ , then for any character  $\chi(b) = \pm 1$ . But,  $\chi_3(b) = -\iota$ , implies that  $|b| = 4$ .

Exactly similar holds in case of  $c$ , and we have  $|c| = 4$ . Exact same argument as the case of  $b$ .

Now, again as in case of  $b$ , we get  $|d| = 2$  or  $4$ . Observe that  $\text{Ker}\chi_2$  is a normal subgroup of  $G$ , and  $d \in \text{Ker}\chi_2$ , with  $|\text{Ker}\chi_2| = 10$ . Therefore  $|d| \mid 10$ , and hence  $|d| = 2$ . This solves (b).

Next, we do (e) first then we go back to solve (c) and (d).

For that we prove a general result which will be of use in many problems of this book. This result actually determines the following: Given a finite group  $G$ , and its character table  $T$ , how can one find all normal subgroups of  $G$ . To do this we first prove the following claim:

**Claim 1** Suppose  $R$  is a representation of  $R : G \rightarrow GL(V)$  of degree  $d$ , and  $\chi_R$  denote the character of  $R$ . Define  $\text{Ker}(\chi_R) = \{g \in G \mid \chi_R(g) = d\}$ . We claim that  $\text{Ker}(\chi_R) = \text{Ker}R$ , where  $\text{Ker}R$  is the usual kernel of the homomorphism  $R$ . Indeed if  $g \in \text{Ker}R$ , then  $R_g = id_V$ . Since degree of  $R$  is  $d$ , we get  $\chi_R(g) = \text{trace}(id_V) = d$ . Therefore  $g \in \text{Ker}(\chi_R)$ . This proves  $\text{Ker}R \subseteq \text{Ker}(\chi_R)$ .

Conversely suppose  $g \in \text{Ker}(\chi_R)$ . This implies  $\chi_R(g) = d \implies |\chi_R(g)| = d$ . Now, it is a standard observation that  $\chi_R(g)$  is sum of  $d$  roots of unity. Also, it is a standard result that if  $\zeta_1, \dots, \zeta_d$  are roots of unity then  $|\zeta_1 + \dots + \zeta_d| \leq d$ , with equality iff  $\zeta_1 = \zeta_2 = \dots = \zeta_d$ . so, from this we conclude that as  $|\chi_R(g)| = d$ , then  $\chi_R = d\zeta$ , where  $\zeta$  is a root of unity. But,  $\chi_R(g) = d$  implies that  $\zeta = 1$ . Hence, we get that all the eigen values of  $R_g$  is 1. But  $R_g$  is diagonalizable (again standard observation, as minimal polynomial of  $R_g$  divides  $X^n - 1$ , where  $n$  is the order of  $g$ ), and therefore we conclude that  $R_g = id_V$ . Therefore,  $\text{Ker}(\chi_R) \subseteq \text{Ker}R$ . This completes the proof of the claim.

Suppose now that  $\chi$  is any character (not necessarily irreducible) of  $G$ . Let  $\chi = \sum_{i=1}^k n_i \chi_i$ , where  $\chi_1, \chi_2, \dots, \chi_k$  denote all the distinct irreducible characters, and  $n_i (1 \leq i \leq k)$  are non-negative integers. Let  $d_i$  denote the degrees of the character  $\chi_i$  for every  $1 \leq i \leq k$ .



**Claim 2:**  $\text{Ker}\chi = \cap \{ \text{Ker}\chi_i | n_i > 0 \}$ .

Let  $A = \cap \{ \text{Ker}\chi_i | n_i > 0 \}$ . Let  $g \in A$ . Then  $\chi_i(g) = d_i$ , for every  $i$  such that  $n_i = 0$ . If  $d$  is the degree of  $\chi$ , it is clear that  $\chi(g) = d$ . Therefore,  $g \in \text{Ker}\chi$ . This proves  $A \subset \text{Ker}\chi$ . Conversely, suppose  $g \in \text{Ker}\chi$ , then  $\chi(g) = d$ . So  $|\chi_i(g)| \leq d_i$ , claim 1 forces  $\chi_i(g) = d_i$ , whenever  $n_i > 0$ . Therefore, we have that  $g \in A$ . Thus, we have proved our second claim.

Now, we proceed towards the final claim, which tells us how to determine all normal subgroups of  $G$ , by the character table of  $G$ . Let,  $N_i = \text{Ker}\chi_i$ . It is clear the  $N_i$  are normal subgroups.

**Claim 3:** Any normal subgroup of  $G$  is a certain intersection of  $N_i$ 's.

Let  $N$  be any normal subgroup of  $G$ . Consider  $G/N$ . Let  $\rho$  denote the regular representation of  $G/N$ . Let  $\pi : G \rightarrow G/N$  be the canonical surjection, with  $\text{Ker}\pi = N$ . Consider  $\rho_G = \rho \circ \pi$ . Observe that since  $\rho$  is injective,  $\rho_G$  is a representation of  $G$ , with  $\text{Ker}\rho_G = N$ . Let  $\chi_G$  be the character of  $\rho_G$ . Then, by claim 1,  $N = \text{Ker}\chi_G$ . Now, Let

$$\chi_G = \sum_{i=1}^k n_i \chi_i$$

. Then, from claim 2, we get  $N = \text{Ker}\chi_G = \cap \{ \text{Ker}\chi_i | n_i > 0 \} = \cap \{ N_i | n_i > 0 \}$ .

So, finally it is easy to determine the  $N_i$ 's from the character table. To determine  $N_i$ , just look at the  $i^{\text{th}}$  row, and find out all elements of  $G$ , whose character entry is  $d_i$ , where  $d_i = \chi_i(1)$ . Then to determine all normal subgroups, just take all possible intersection of these  $N_i$ 's, and list the new subgroups thus obtained. This gives you all the possible normal subgroups.

So,

$$\begin{aligned} N_1 &= \{g \in G | \chi_1(g) = 1\} = G \\ N_2 &= \{g \in G | \chi_2(g) = 1\} = \{e\} \cup Cl(a) \cup Cl(d) \\ N_3 &= \{g \in G | \chi_3(g) = 1\} = \{e\} \cup Cl(a) \\ N_4 &= \{g \in G | \chi_4(g) = 1\} = \{e\} \cup Cl(a) \\ N_5 &= \{g \in G | \chi_5(g) = 4\} = \{e\} \end{aligned}$$

Therefore to get other normal subgroup, we take all possible intersection of these  $N_i$ 's according to claim 3. Observe that there are no other are new normal subgroup occurring out of intersections. Hence  $N_i (1 \leq i \leq 5)$  are the only normal subgroups of  $G$ . This solves (e) as well as (c), as we can take  $H = N_2$ . Therefore,  $H = \{e\} \cup Cl(a) \cup Cl(d)$

Now, we have to understand to what is  $H$  isomorphic to. We know that since  $|H| = 10$ ,  $H \cong C_{10}$  or  $D_5$ . Now,  $a, d \in H$ , and  $|a| = 5, |d| = 2$ . Suppose  $H \cong C_{10}$ , then there exists  $x \in H$  such that  $|x| = 10$ . But, since we have seen that  $|a| = 5, |d| = 2, |b| = 4 = |c|$ , it is clear that any element of  $G$  has order 1, 2, 4, or 5. Therefore,  $H \cong D_5$ . This solves (d).

## Result

6 of 6

We have proved that  $H \cong D_5$ . To determine all normal subgroup of  $G$  from the character table, we have first explained a general method of doing so, and then applied the general case in our situation. See the solution for more details.



Let  $\omega = e^{2\pi i/3}$ . Consider the following character table.

	(1)	(6)	(7)	(7)	(7)	(7)	(7)
	1	a	b	c	d	e	f
$\chi_1$	1	1	1	1	1	1	1
$\chi_2$	1	1	1	$\omega$	$\bar{\omega}$	$\omega$	$\bar{\omega}$
$\chi_3$	1	1	1	$\bar{\omega}$	$\omega$	$\bar{\omega}$	$\omega$
$\chi_4$	1	1	-1	$-\omega$	$-\bar{\omega}$	$\omega$	$\bar{\omega}$
$\chi_5$	1	1	-1	$-\bar{\omega}$	$-\omega$	$\bar{\omega}$	$\omega$
$\chi_6$	1	1	-1	-1	-1	1	1
$\chi_7$	6	-1	0	0	0	0	0

First counting the size of the conjugacy classes we see that  $|G| = 42$ .

(a) We have to show that there is a normal subgroup  $N$  of  $G$ , such that  $G$  is isomorphic  $D_7$ . Observe that  $\text{Ker}\chi_3$  is a normal subgroup and  $N = \text{Ker}\chi_3 = 1 \cup Cl(a) \cup Cl(b)$ , where  $Cl(x)$  denote the conjugacy class of  $x \in G$ . Then, it is clear that  $|N| = 14$ . We claim that  $N \cong D_7$ . To prove this observe that  $K$  has elements of order  $m$  and  $n$  only, where  $o(a) = m, o(b) = n$ . So, all non-trivial elements of  $K$  has only two possible choices of order. Now, if  $K$  is cyclic of order 14, and hence it has elements of order 2, 7, 14. This is a contradiction. Therefore,  $N$  is cyclic. Since every group of order 14, is either the dihedral group, or cyclic group of order 14, we conclude that  $N \cong D_7$ .

Now, first let us write  $N \cong D_7$  in the proper form. Observe that  $C(a)$  has order 7, as  $|Cl(a)| = 6$ . Here  $C(a)$  denote the centralizer of  $a \in G$ . So  $C(a)$  is cyclic, since its order is prime, and therefore each non trivial element of  $C(a)$  has order 7. Therefore, we conclude that  $|a| = 7$ . Therefore, each element of  $Cl(a)$  has order 7. Now,  $D_7$  abstractly as a group, is

$$D_7 = \langle x, y | x^7 = 1 = y^2, xy = y^{-1}x \rangle$$

. So,  $D_7 = \{1, x, x^2, x^3, x^4, x^5, x^6, y, yx, yx^2, yx^3, yx^4, yx^5, yx^6\}$ . Their conjugacy class is well known is as follows:

$$\{1\}, \{x, x^6\}, \{x^2, x^5\}, \{x^3, x^4\}, \{y, yx, yx^2, yx^3, yx^4, yx^5, yx^6\}$$

So since,  $N \cong D_7$ , we observe that  $|b| = 2$ , and hence each element in  $Cl(b)$  must have order 2. Hence,  $N = \langle a, b | a^7 = 1 = b^2, ab = b^{-1}a \rangle$ . The conjugacy classes of  $N$  are therefore, exactly same as  $D_7$ , where  $x$  is replaced by  $a$ , and  $y$  by  $b$ . Therefore, the conjugacy class of  $N$  are

$$C_1 = \{1\}, C_2 = \{a, a^6\}, C_3 = \{a^2, a^5\}$$

,

$$C_4 = \{a^3, a^4\}, C_5 = \{b, ba, ba^2, ba^3, ba^4, ba^5, ba^6\} = Cl(b)$$

Now, we can solve (b). Observe that the restriction of  $\chi_1, \chi_2, \chi_3$  to  $N$  is trivial and hence is the trivial irreducible character of  $N$ .

Next take the character  $\chi_4$ , which is also of degree 1. Then

$$\chi_4(C_1) = 1, \chi_4(C_2) = 1, \chi_4(C_3) = 1, \chi_4(C_4) = 1, \chi_4(C_5) = -1$$

Hence,  $\chi_4$  on  $N$  is the "sign" (degree one) representation of  $N = D_7$ . The same case also happens for  $\chi_5, \chi_6$ .

Next we try to understand the restriction of  $\chi_7$  on  $D_7$ . Observe that

$$\chi_7(C_1) = 6, \chi_7(C_2) = -1, \chi_7(C_3) = -1, \chi_7(C_4) = -1, \chi_7(C_5) = 0$$

To understand how the above representation of  $N$  decomposes into irreducible representation of  $N$ , we need to understand the irreducible representations of  $N$ . We already know that  $N$  has two irreducible degree one representations namely, the trivial one, and the sign representation. Let their characters be  $\theta_1, \theta_2$  respectively. Let the other characters be  $\theta_i (3 \leq i \leq 5)$ . There are 5 irreducible representations because there are 5 conjugacy classes. Let their degrees be  $d_i (1 \leq i \leq 5)$ . It is clear that

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 = 14$$

$d_1 = d_2 = 1$ . Now, we say that the only way  $d_3^2 + d_4^2 + d_5^2 = 12$  is if each of  $d_3 = d_4 = d_5 = 2$ . Hence, it follows that  $N$  has three irreducible representation of degree 2. Now, I will write the character table of  $N \cong D_7$ .

	(1)	(2)	(2)	(2)	(7)
	1	$a$	$a^2$	$a^3$	$b$
$\theta_1$	1	1	1	1	1
$\theta_2$	1	1	1	1	-1
$\theta_3$	2	$2 \cos 2\pi/7$	$2 \cos 4\pi/7$	$2 \cos 6\pi/7$	0
$\theta_4$	2	$2 \cos 4\pi/7$	$2 \cos 6\pi/7$	$2 \cos 2\pi/7$	0
$\theta_5$	2	$2 \cos 6\pi/7$	$2 \cos 2\pi/7$	$2 \cos 4\pi/7$	0

I mention here that I leave it to the reader to verify that  $\theta_3, \theta_4, \theta_5$  are indeed the remaining three irreducible characters of  $N$ . The proof of this will be exactly the same as for the construction of the character table of  $D_5$ , which I have done myself, in the problem 4 of section 4. Mimic the exact same ideas to construct the character table of  $N \cong D_7$ .

Now, observe that

$$\langle \chi_7^N, \theta_3 \rangle = \frac{1}{14} [6 \cdot 2 + -2 \cdot 2 \cdot (\cos 2\pi/7) + -2 \cdot 2 \cdot (\cos 4\pi/7) + -2 \cdot 2 \cdot (\cos 6\pi/7) + 7 \cdot 0 \cdot 0]$$

=

$$\frac{1}{14} [12 - 4((\cos 2\pi/7) + (\cos 4\pi/7) + (\cos 6\pi/7))] = 1$$

.

This is because  $\cos 2\pi/7 + \cos 4\pi/7 + \cos 6\pi/7 = -\frac{1}{2}$ . Similarly,

$$\langle \chi_7^N, \theta_4 \rangle = \langle \chi_7^N, \theta_5 \rangle = 1$$

. So we finally get,

$$\chi_7^N = \theta_3 + \theta_4 + \theta_5$$

This finishes (b).

Before doing (c) and (d), we will first do e. We will use the following general principle. This result actually determines the following: Given a finite group  $G$ , and its character table  $T$ , how can one find all normal subgroups of  $G$ . To do this we first prove the following claim:

**Claim 1** Suppose  $R$  is a representation of  $R : G \rightarrow GL(V)$  of degree  $d$ , and  $\chi_R$  denote the character of  $R$ . Define  $Ker(\chi_R) = \{g \in G | \chi_R(g) = d\}$ . We claim that  $Ker(\chi_R) = Ker R$ , where  $Ker R$  is the usual kernel of the homomorphism  $R$ . Indeed if  $g \in Ker R$ , then  $R_g = id_V$ . Since degree of  $R$  is  $d$ , we get  $\chi_R(g) = \text{trace}(id_V) = d$ . Therefore  $g \in Ker(\chi_R)$ . This proves  $Ker R \subseteq Ker(\chi_R)$ .

Conversely suppose  $g \in Ker(\chi_R)$ . This implies  $\chi_R(g) = d \implies |\chi_R(g)| = d$ . Now, it is a standard observation that  $\chi_R(g)$  is sum of  $d$  roots of unity. Also, it is a standard result that if  $\zeta_1, \dots, \zeta_d$  are roots of unity then  $|\zeta_1 + \dots + \zeta_d| \leq d$ , with equality iff  $\zeta_1 = \zeta_2 = \dots = \zeta_d$ . so, from this we conclude that as  $|\chi_R(g)| = d$ , then  $\chi_R = d\zeta$ , where  $\zeta$  is a root of unity. But,  $\chi_R(g) = d$  implies that  $\zeta = 1$ . Hence, we get that all the eigen values of  $R_g$  is 1. But  $R_g$  is diagonalizable (again standard observation, as minimal polynomial of  $R_g$  divides  $X^n - 1$ , where  $n$  is the order of  $g$ ), and therefore we conclude that  $R_g = id_V$ . Therefore,  $Ker(\chi_R) \subseteq Ker R$ . This completes the proof of the claim.

Suppose now that  $\chi$  is any character (not necessarily irreducible) of  $G$ . Let  $\chi = \sum_{i=1}^k n_i \chi_i$ , where  $\chi_1, \chi_2, \dots, \chi_k$  denote all the distinct irreducible characters, and  $n_i (1 \leq i \leq k)$  are non-negative integers. Let  $d_i$  denote the degrees of the character  $\chi_i$  for every  $1 \leq i \leq k$ .

**Claim 2:**  $Ker \chi = \cap \{ Ker \chi_i | n_i > 0 \}$ .

Let  $A = \cap \{ Ker \chi_i | n_i > 0 \}$ . Let  $g \in A$ . Then  $\chi_i(g) = d_i$ , for every  $i$  such that  $n_i > 0$ . If  $d$  is the degree of  $\chi$ , it is clear that  $\chi(g) = d$ . Therefore,  $g \in Ker \chi$ . This proves  $A \subseteq Ker \chi$ . Conversely, suppose  $g \in Ker \chi$ , then  $\chi(g) = d$ . So  $|\chi_i(g)| \leq d_i$ , claim 1 forces  $\chi_i(g) = d_i$ , whenever  $n_i > 0$ . Therefore, we have that  $g \in A$ . Thus, we have proved our second claim.

Now, we proceed towards the final claim, which tells us how to determine all normal subgroups of  $G$ , by the character table of  $G$ . Let,  $N_i = Ker \chi_i$ . It is clear the  $N_i$  are normal subgroups.

**Claim 3:** Any normal subgroup of  $G$  is a certain intersection of  $N_i$ 's.

Let  $N$  be any normal subgroup of  $G$ . Consider  $G/N$ . Let  $\rho$  denote the regular representation of  $G/N$ . Let  $\pi : G \rightarrow G/N$  be the canonical surjection, with  $Ker \pi = N$ . Consider  $\rho_G = \rho \circ \pi$ . Observe that since  $\rho$  is injective,  $\rho_G$  is a representation of  $G$ , with  $Ker \rho_G = N$ . Let  $\chi_G$  be the character of  $\rho_G$ . Then, by claim 1,  $N = Ker \chi_G$ . Now, Let

$$\chi_G = \sum_{i=1}^k n_i \chi_i$$

. Then, from claim 2, we get  $N = Ker \chi_G = \cap \{ Ker \chi_i | n_i > 0 \} = \cap \{ N_i | n_i > 0 \}$ .

So, finally it is easy to determine the  $N_i$ 's from the character table. To determine  $N_i$ , just look at the  $i^{th}$  row, and find out all elements of  $G$ , whose character entry is  $d_i$ , where  $d_i = \chi_i(1)$ . Then to determine all normal subgroups, just take all possible intersection of these  $N_i$ 's, and list the new subgroups thus obtained. This gives you all the possible normal subgroups.

Therefore,

$$\begin{aligned} N_1 &= \text{Ker}\chi_1 = G, N_2 = \text{Ker}\chi_2 = \{1\} \cup \text{Cl}(a) \cup \text{Cl}(b) = N, N_3 = \text{Ker}\chi_3 = N_2 = N \\ N_4 &= \text{Ker}\chi_4 = \{1\} \cup \text{Cl}(a), N_5 = \text{Ker}\chi_5 = N_4, N_6 = \text{Ker}\chi_6 = \{1\} \cup \text{Cl}(a) \cup \text{Cl}(e) \cup \text{Cl}(f) \\ N_7 &= \text{Ker}\chi_7 = \{1\} \end{aligned}$$

So, we have got five distinct normal subgroups namely  $G, N, N_4, N_6, \{1\}$ . Now, using the general principle we observe that any other normal subgroup is intersection of these subgroups. We observe that we don't get any new normal subgroup, by intersecting them with each other in any way. So, we conclude that these are the only five normal subgroups of  $G$ . For clarity  $|N| = 14, |N_4| = 7, |N_6| = 21$ .

Now, we have to determine the Sylow- $p$  subgroups of  $G$ . We will use the description of normal subgroups to our advantage.  $|G| = 42 = 2 \cdot 3 \cdot 7$ . First observe that the sylow-7 subgroup must have order 7. Now,  $N_4$  is a normal subgroup of  $G$  of order 7. We know that there is a unique sylow- $p$  subgroup of order  $p$  if and only if the subgroup is normal. This is precisely the Sylow second theorem. Therefore we conclude that there is a unique Sylow-7 subgroup of order 7, which is  $N_4$ .

Next we try to determine the number of Sylow-3 subgroup which are of order 3. Let  $n_3$  be the number of Sylow-3 subgroups. Then  $n_3 \mid 14$ . So  $n_3 = 1$  or 7. Now, observe that if  $n_3 = 1$ , then that means there is a unique normal subgroup of order 3, and hence it must be normal. But, we see that in our complete list of normal subgroups there is such subgroup of order 3. Therefore, we conclude, that  $n_3 = 7$ , and hence there are 7 Sylow-3 subgroups.

Finally, we determine the number of Sylow-2 subgroups. Again any Sylow-2 subgroups has order 2. Let  $n_2$  be the number of Sylow-2 subgroups. Then  $n_2 \mid 21$ , which means  $n_2 = 1$  or 3 or 7 or, 21. Now, again by Sylow second theorem, and from our complete list of normal subgroups, we conclude that  $n_2 \neq 1$ . Also, observe that we have already seen that each element of conjugacy class of  $\text{Cl}(b)$  has order 2. So, we already have 6 sylow-2 subgroups namely  $\{1, b\}$  and its six distinct conjugates. So  $n_2 \neq 3$ . so, now we have  $n_2 = 7$  or 21.

Now observe also that if  $n_2 = 7$ , there are 7 elements of order 2, and if  $n_2 = 21$ , there are 21 elements of order 2. Now, we now know that there are 7 elements of order 2, that is, elements of  $\text{Cl}(b)$ .

Now,  $\chi_2(c) = \omega$ , implies that  $3 \mid |c|$ , as order of  $\omega$  in  $\mathbb{C}^*$  is 3. So, we conclude that no element of  $\text{Cl}(c)$  has order 2. Similarly, we conclude that no element of  $\text{Cl}(d), \text{Cl}(e), \text{Cl}(f)$  has order 2. Therefore there are only 7 elements of order 2, which forces  $n_2$  to be 7.

So, to summarize, there is exactly one Sylow-7 subgroup, 7 Sylow-3 subgroup and 21 Sylow 2-subgroup. This completes (c).

Now, we finally answer (d) and finish the problem. Observe that  $e, f \in N_6$ , and  $|N_6| = 21$ . So,  $|e|, |f|$  divides 21. Now, also observe that since  $|\text{Cl}(e)| = |\text{Cl}(f)| = 7$ , we conclude from the orbit-stabilizer theorem, that  $|C(e)| = |C(f)| = 6$ . Since  $e \in C(e)$ , we conclude that  $|e| \mid 6$ . Similarly,  $|f| \mid 6$ . So, we have that  $|e| = |f| = 3$ . Now, again for the same reason as above we have that  $|c|, |d|$  divides 6. In the discussion of the Sylow-2 subgroups above we have already concluded that  $|c|, |d| \neq 2$ . So, we have  $|c|$  is either 3 or 6. Similarly,  $|d|$  is either 3 or 6. But, since  $|e| = |f| = 3$ , so each element of conjugacy class of  $e$  and  $f$  has order 3. So, we get 14 elements of order 3. Since the number of Sylow-3 subgroup is 7, we have that  $G$  has 14 elements of order 3. So, conjugacy class of  $e$  and  $f$  account for all the elements of order 3, and therefore  $|c| \neq 3, |d| \neq 3$ . Therefore, we conclude that  $|c| = |d| = 6$ . This completes the proof.

## Result

10 of 10

There is exactly one Sylow-7 subgroup, 7 Sylow-3 subgroup and 21 Sylow 2-subgroup of  $G$ . Also  $|c| = |d| = 6, |e| = |f| = 3$ . There are five distinct normal subgroups of  $G$ , which has been clearly described in the answer.

See the answer for more detailed explanations.



Let  $H$  be a subgroup of index 2 in  $G$ . Suppose  $\sigma : H \rightarrow GL(V)$ , be a representation. Let  $a \in G \setminus H$ . Define a representation  $\sigma' : H \rightarrow GL(V)$ , defined by  $\sigma'(h) = \sigma(a^{-1}ha)$ , for every  $h \in H$ . First we prove that  $\sigma'$  is a representation of  $H$ . Let  $h_1, h_2 \in H$ . Then  $\sigma'(h_1h_2) = \sigma(a^{-1}h_1h_2a) = \sigma(a^{-1}h_1aa^{-1}h_2a) = \sigma(a^{-1}h_1a)\sigma(a^{-1}h_2a) = \sigma'(h_1)\sigma'(h_2)$ . This prove that  $\sigma'$  is a representation.

## Step 2

2 of 4

Next, suppose  $\sigma$ , is a restriction of a representation say,  $\theta$  of  $G$ . We want to prove that  $\sigma$ , and  $\sigma'$ , are equivalent representation. Let  $\chi, \chi'$ , denote the characters of  $\sigma$ , and  $\sigma'$  respectively. In fact, we can assume  $\chi$  to be the character of  $\theta$  of  $G$ , and without loss of generality, we denote the restriction of  $\chi$  to  $H$  by  $\chi$  again. With these notations in hand, we will prove  $\chi = \chi'$ , on  $H$ . Now, by definition  $\chi(h) = \text{trace}(\sigma(h)) = \text{trace}(\theta(h))$ . But  $\chi'(h) = \text{trace}(\sigma(a^{-1}ha)) = \text{trace}(\theta(a^{-1}ha)) = \text{trace}(\theta(a)^{-1}\theta(h)\theta(a)) = \text{trace}(\theta(h)) = \text{trace}(\sigma(h)) = \chi(h)$ . The previous equalities follow from the fact that  $\theta$  is defined on  $G$ . So, we get  $\chi' = \chi$ , and hence  $\sigma, \sigma'$  are isomorphic.

Finally, let  $b \in G \setminus H$ , and  $b \neq a$ . Define  $\sigma'' : H \rightarrow GL(V)$ , as  $\sigma''(h) = \sigma(b^{-1}hb)$ . We have to prove that  $\sigma$ , and  $\sigma''$  are isomorphic. Let  $\chi''$ , be the character of  $\sigma''$ . We again wish show that  $\chi' = \chi''$ . The major fact that we use here is that since  $[G : H] = 2$ , we have that  $Ha = aH = Hb = bH = G \setminus H$ . Also, we use the fact that  $\text{trace}(AB) = \text{trace}(BA)$ . Now, observe that  $\chi'(h) = \text{trace}(\sigma(a^{-1}ha))$ . Now, as  $Ha = bH$ , we have that  $ha = bh_1$ , for some  $h_1 \in H$ . So,  $\text{trace}(\sigma(a^{-1}ha)) = \text{trace}(\sigma(a^{-1}bh_1)) = \text{tr}(\sigma(a^{-1}b)\sigma(h_1))$ . The last equality is possible because  $a^{-1}b \in H$ . From there we have that  $\text{trace}(\sigma(a^{-1}b)\sigma(h_1)) = \text{trace}(\sigma(h_1)\sigma(a^{-1}b)) = \text{trace}(\sigma(h_1a^{-1}b)) = \text{trace}(\sigma(b^{-1}hb)) = \chi''(h)$ . The last equality follows, because,  $ha = bh_1 \implies h_1a^{-1} = b^{-1}h$ . So, finally we have  $\chi' = \chi''$ , and therefore  $\sigma', \sigma''$ , are isomorphic.

## Result

4 of 4

We have proved that the representations are isomorphic by showing that their characters are equal. While showing that there characters are equal, we have heavily used the fact that  $[G : H] = 2$ . See the solution for more details.

# Section 5

1. a

Let  $C_n$  be the cyclic group of order  $n$ . Suppose  $g$  is a generator of  $C_n$ . The two dimensional standard representation of  $C_n$ , can be given as follows:

$\phi : C_n \rightarrow GL_2(\mathbb{C})$ , defined by,

$$\phi(g^i) = \begin{bmatrix} \cos \frac{2\pi i}{n} & \sin \frac{2\pi i}{n} \\ -\sin \frac{2\pi i}{n} & \cos \frac{2\pi i}{n} \end{bmatrix}$$

for every  $1 \leq i \leq n-1$ .

Let  $\chi$  be the character of  $\phi$ . Then,  $\chi(g^i) = 2 \cos \frac{2\pi i}{n} = e^{\frac{2\pi i}{n}} + e^{-\frac{2\pi i}{n}}$ . We have to decompose  $\phi$ , into two irreducible representation of dimension 1. Consider the one-dimensional representations  $\chi_1, \chi_2$ , defined as,

$\chi_1 : C_n \rightarrow \mathbb{C}^*$ , as  $\chi_1(g^i) = e^{\frac{2\pi i}{n}}$ , and  $\chi_2 : G \rightarrow \mathbb{C}^*$ , as  $\chi_2(g^i) = e^{-\frac{2\pi i}{n}}$ . It is clearly seen that  $\chi_1, \chi_2$ , are representation of dimension 1. We can clearly see that  $\chi = \chi_1 + \chi_2$ , and therefore, we have that  $\phi$  is the direct sum of two one dimensional representations  $\chi_1$ , and  $\chi_2$ .

## Result

2 of 2

If  $\chi$  the character of the standard representation of  $C_n$ , we have proved  $\chi = \chi_1 + \chi_2$ , where,  $\chi_1 : C_n \rightarrow \mathbb{C}^*$ , is defined as  $\chi_1(g^i) = e^{\frac{2\pi i}{n}}$ , and  $\chi_2 : G \rightarrow \mathbb{C}^*$ , as  $\chi_2(g^i) = e^{-\frac{2\pi i}{n}}$

## 2. a

Consider  $G = S_n$ ,  $n \geq 2$ . We will have to prove that  $S_n$  has only two one-dimensional representation, the trivial representation and the sign representation,  $p \mapsto \text{sign}(p)$ .

For  $n = 2$ ,  $S_2$  is cyclic group of order 2, and hence there are only two representations as above. So, from now on we assume  $n \geq 3$ .

Suppose  $\phi : S_n \rightarrow \mathbb{C}^*$ , be any representation. Now, it is clear that  $S'_n \subseteq \text{Ker} \phi$ , where  $S'_n$ , denote the derived subgroup of  $S_n$ . We claim that  $S'_n = A_n$ . For the time being assuming the claim, we observe that  $\text{Ker} \phi$ , is either  $S_n$ , or  $A_n$ . Suppose  $\text{Ker} \phi$  is  $S_n$ , then it is clear that  $\phi$  is trivial representation. On the other hand if  $\text{Ker} \phi$  is  $A_n$ , then  $\phi$  induces the map  $pA_n \mapsto \text{to}\phi(p)$ , from  $S_n/A_n \rightarrow \mathbb{C}^*$ . Since  $S_n/A_n = \{A_n, pA_n\}$ , where  $p$  is any odd permutation, is a cyclic group of order 2, it is clear that  $\phi(p) = -1$ . Therefore, we conclude that if  $\text{Ker} \phi = A_n$ , then  $\phi$  is the sign map. Hence, there are only two one-dimensional representation of  $S_n$ .



Now, we prove the claim that  $S'_n = A_n$ , for  $n \geq 3$ . Since  $S_n/A_n$  is of size 2, and hence abelian, we conclude that  $S'_n \subseteq A_n$ . Now, we prove that any 3-cycle is a commutator. Let  $(abc)$  be a 3-cycle. Observe that  $(ab)(ac)(ab)^{-1}(ac)^{-1} = (ab)(ac)(ab)(ac) = (abc)$ . So, indeed any 3-cycle is a commutator.

Next we prove that  $A_n$  is generated by 3-cycles. First observe that all 3-cycles are contained in  $A_n$ , since  $(abc) = (ab)(ac)$ . For  $n = 3$ ,  $A_3$  consists only of 3-cycles, and hence there is nothing to prove. So, we assume that  $n \geq 4$ . We know that if  $p \in A_n$ , then  $p$  is a product of even number of transposition. So, it is enough to prove that a product of two transposition is a product of 3-cycle. So, suppose if there are two transposition, with one element in common, say  $(ab)$ , and  $(ac)$ , then  $(ab)(ac) = (acb)$ . Hence, their product is a 3-cycle. Now, assume there are two transposition with distinct numbers, that is,  $(ab)$ , and  $(cd)$ , then  $(ab)(cd) = (dac)(abd)$ . So, in this case also, we see that their product is a product of 3-cycles. So, any product of two transposition is a product of 3-cycles. Therefore, any permutation is  $A_n$ , being a product of even number of transposition, is also a product of 3-cycles. Hence,  $A_n$  is generated by 3-cycle.

With this, now we see that since every 3-cycle is a commutator, we conclude that  $A_n \subseteq S'_n$ . So, we get  $S'_n = A_n$ , thereby completing the proof of the claim.

### Result

3 of 3

The main result that we use is that the commutator subgroup of  $S_n$  is  $A_n$ . See the solution for more details.

### 3. a

Let  $G$  be a finite group, which has exactly two irreducible characters of degree 1. Suppose  $\phi$  denote the trivial one, that is,  $\phi(g) = 1$  for every  $g \in G$ . Suppose  $\chi$  denote the non-trivial character. Since  $\chi$  is a character of degree 1, hence  $\chi$  is a homomorphism from  $G$  to  $\mathbb{C}^*$ . Define the following map  $\chi^2 : G \rightarrow \mathbb{C}^*$  defined by  $\chi^2(g) = \chi(g)^2$ . It is easy to check that  $\chi^2$  is a group homomorphism. Indeed,  $\chi^2(gh) = \chi(gh)^2 = (\chi(g)\chi(h))^2 = \chi(g)^2\chi(h)^2 = \chi^2(g)\chi^2(h)$ , for every  $g, h \in G$ . Therefore,  $\chi^2$  is a character of  $G$  of degree 1. Hence two possiblity arise:

- (1)  $\chi^2 = \phi$ . Then  $\chi(g)^2 = 1$ , for every  $g \in G$ , which implies  $\chi(g) = \pm 1$ , and we are done.
- (2) Suppose  $\chi^2 = \chi$ , whence, we have  $\chi(g)^2 = \chi(g) \implies \chi(g)(\chi(g) - 1) = 0$ , for every  $g \in G$ . We conclude that  $\chi(g) = 1$ , for every  $g \in G$ . This is a contradiction, as  $\chi$  is non-trivial. Therefore only (1) can happen, whence,  $\chi(g) = \pm 1$ , for every  $g \in G$ . This finishes the proof.

### Result

2 of 2

The main idea of the proof is to observe that  $\chi^2 : G \rightarrow \mathbb{C}^*$ , defined by  $\chi^2(g) = \chi(g)^2$ , is a group homomorphism. This along with the hypothesis of the problem yields the proof. See the solution for more details.

### 4. a

Let  $\chi$  be a character of a finite group  $G$ , say of order  $n$ . Let  $g \in G$ , and assume  $\rho : G \rightarrow GL(V)$  is the representation of  $G$  affording  $\chi$ . Now, observe that  $g^n = 1$ , and therefore, since  $\rho$  is a homomorphism of groups  $\rho(g)^n = I$ . This shows that the polynomial  $X^n - 1$  annihilates  $\rho(g)$ , and therefore if  $M$ , denote its minimal polynomial then  $M \mid X^n - 1$ . But, the minimal polynomial is the product of distinct eigen values, and therefore, we conclude that the eigen values are roots of unity. Now, by definition  $\chi(g)$ , is the trace of  $\rho(g)$ , and we know that the trace of a matrix is the sum of eigen values (with multiplicity). Therefore, we see that  $\chi(g)$  is the sum of  $d$  roots of unity say  $\zeta_1, \zeta_2 \dots \zeta_d$  (not necessarily distinct). Therefore

$$|\chi(g)| = |\zeta_1 + \dots \zeta_d| \leq |\zeta_1| + \dots + |\zeta_d| = d.$$

So, the first part of the problem has been proved.

Now, we prove that if  $|\chi(g)| = d$ , then,  $\rho(g) = \zeta I$ , where  $\zeta$ , is some root of unity. We use the following standard fact:

Suppose  $\zeta_1, \zeta_2 \dots \zeta_d$  are roots of unity. Then, we have  $|\zeta_1 + \zeta_2 + \dots \zeta_d| = d$ , iff  $\zeta_1 = \zeta_2 = \dots = \zeta_d$ .

Now,  $|\chi(g)| = d$ , implies that  $|\zeta_1 + \zeta_2 + \dots \zeta_d| = d$ , and therefore by the above fact,  $\zeta_1 = \zeta_2 = \dots = \zeta_d = \zeta$ . Now, another point to observe is the fact that each  $\rho(g)$  is a diagonalizable operator, or in other words their minimal polynomial have distinct roots. This is clear because the minimal polynomial divides  $X^n - 1$ , which itself has distinct roots. Now, from above, we see that all eigen values are same, which is  $\zeta$ , and therefore,  $\rho(g) = \zeta I$ .

Now, suppose  $\chi(g) = d$ , then  $|\chi(g)| = d$ , and therefore from above, we have  $\rho(g) = \zeta I$ , for some root of unity  $\zeta$ . Therefore,  $\chi(g) = d\zeta = d$ , and hence,  $\zeta = 1$ . Therefore,  $\rho(g) = I$ .

## Result

2 of 2

The important fact to observe for this problem is that if  $\chi$  is a character of  $G$ , then  $\chi(g)$  is a sum of roots of unity. See the proof for more details.

## 5. a

To prove that the one-dimensional characters of a group  $G$  form a group under multiplication of functions, that is, where the group operation is:

$$(\chi \cdot \chi')(g) = \chi(g)\chi'(g)$$

This prove will be show if the product of any two non-zero real numbers is itself a non-zero real number, so the set is closed under multiplication.

Let  $g \in G$ , and  $\chi, \chi'$  are a one-dimensional characters of a group  $G$ .

$$\begin{aligned} (\chi \cdot \chi')(g) &= \rho_g \rho'_g \\ &= \chi(g)\chi'(g) \end{aligned}$$

Hence, the one-dimensional characters of a group  $G$  form a group under multiplication of functions,

To show that  $\hat{G} = G$  and  $|\hat{G}| = |G|$  if  $G$  is abelian.

Since a one-dimensional character  $\chi$  is a homomorphisms from  $G$  to  $GL_1 = \mathbb{C}^*$ , because

$$\begin{aligned} (\chi)(gh) &= \rho_{gh} \\ &= \rho_g \rho_h \\ &= \chi(g)\chi(h) \end{aligned}$$

Where,

$$g, h \in G$$

$\hat{G}$  is abelian regardless of  $G$  being abelian, since  $\hat{G} \subset \mathbb{C}^*$  and  $\mathbb{C}^*$  is commutative.

So,  $G$  is abelian

$$(\chi \cdot \chi')(gh) = (\chi \cdot \chi')(hg) \text{ With } gh \neq hg$$

To show that  $G$  and  $\hat{G}$  are isomorphic.

A character of a finite abelian group  $G$  is a homomorphism  $\chi: G \rightarrow S^1$  where  $\hat{G} = S^1$  because  $(\chi)(gh) = \chi(g)\chi(h)$  and  $\chi(1) = 1$ .

So,  $G$  and  $\hat{G}$  are isomorphic, that is  $G \cong \hat{G}$ .

So, number of the element in group  $G$  and  $\hat{G}$  have the same,  $\hat{G} \sim G$ .

It can be written as cardinality of  $G$  and  $\hat{G}$  have the same,  $|\hat{G}| = |G|$ .

**Hence proved**

6. a

Let  $G$  be a cyclic group of order  $n$ , generated by an element  $x$ , and let  $\zeta = e^{\frac{2\pi i}{n}}$ .

[Comment](#)

Step 2 of 3 ^

(a)

A representation  $\rho_k: G \rightarrow \mathbb{C}^*$  is defined by,

$$\rho_k(x) = \zeta^k$$

To show that the irreducible representation are  $\rho_0, \dots, \rho_{n-1}$ ;

Since  $\rho_0, \dots, \rho_{n-1}$  are all one dimensional representation and it is known that every irreducible character of  $G$  is one-dimensional.

Thus, it is irreducible.

Moreover, since  $G$  is cyclic of order  $n$ , that is  $|G| = n$  and also every conjugacy class of group  $G$  has size 1.

So, there are  $n$  conjugacy classes of  $G$  and  $n$  irreducible representations.

**Hence,**  $\rho_0, \dots, \rho_{n-1}$  are the irreducible representations of  $G$ .

(b)

To identify the character group of  $G$ ;

The character group of  $G$  is abelian, so apply the one-dimensional characters of a  $G$  form a group under multiplication of function,  $G$  is abelian then  $|\hat{G}| = |G|$  and  $\hat{G} \cong G$ .

So, this implies that  $\boxed{G \cong Z_n}$  where  $Z_n$  is group of order  $n$  and it is abelian and cyclic.

7. a

**Given:**  $G$  and  $G'$  are abelian groups, and  $\phi : G \rightarrow G'$  is a homomorphism.

**Solution:** First we define an induced homomorphism

$$\hat{\phi} : \hat{G}' \rightarrow \hat{G}$$

where  $\hat{G}$  and  $\hat{G}'$  are respective Character groups.

Let us define  $\hat{\phi}$  by the assignment

$$\hat{\phi}(\chi) = \chi \circ \phi, \quad \text{where } \chi \in \hat{G}'.$$

Clearly it is an homomorphism and satisfies the required conditions.

Now we will show that if  $\hat{\phi}$  is surjective then  $\phi$  is injective and vice-versa.

Let us now assume that  $\hat{\phi}$  is surjective.

Let us now consider two distinct elements from  $G$  as  $g, h$  such that

$$\phi(g) = \phi(h).$$

Let  $\psi \in \hat{G}$ .

Since  $\hat{\phi}$  is surjective there exists  $\chi \in \hat{G}'$  such that

$$\hat{\phi}(\chi) = \psi.$$

This follows that

$$\chi \circ \phi = \psi.$$

Now notice that

$$\begin{aligned} \phi(g) = \phi(h) &\implies (\chi \circ \phi)(g) = (\chi \circ \phi)(h) \\ &\implies \chi(\phi(g)) = \chi(\phi(h)) \\ &\implies \psi(g) = \psi(h), \quad \text{by the definition of } \hat{\phi} \\ &\implies g = h, \quad \text{since } \psi \text{ is a character of } G. \end{aligned}$$

But this is a contradiction, since we have taken  $g$  and  $h$  to be distinct.

Therefore  $\phi$  is injective clearly.

Now we will propose to prove that  $\hat{\phi}$  is injective if  $\phi$  is surjective.

To show the injectivity of  $\hat{\phi}$  let us assume  $\chi_1, \chi_2 \in \hat{G}'$  such that

$$\hat{\phi}(\chi_1) = \hat{\phi}(\chi_2).$$

We will assert that  $\chi_1 = \chi_2$ . That is

$$\chi_1(g') = \chi_2(g') \quad \forall g' \in G'$$

Let  $g'$  be an arbitrary element of  $G'$ . Since  $\phi$  is surjective there exists an element  $g \in G$  such that

$$\phi(g) = g'.$$

Now notice that

$$\begin{aligned} \hat{\phi}(\chi_1) = \hat{\phi}(\chi_2) &\implies \chi_1 \circ \phi = \chi_2 \circ \phi \\ &\implies (\chi_1 \circ \phi)(g) = (\chi_2 \circ \phi)(g) \\ &\implies \chi_1(\phi(g)) = \chi_2(\phi(g)) \\ &\implies \chi_1(g') = \chi_2(g'). \end{aligned}$$

Since  $g'$  is arbitrary and  $\chi_1(g') = \chi_2(g')$  it yields that

$$\chi_1(g') = \chi_2(g') \quad \text{for all } g' \in G'.$$

Therefore

$$\chi_1 = \chi_2.$$



Hence  $\hat{\phi}$  is injective.  
This completes the proof.

## Result

3 of 3

The induced homomorphism  $\hat{\phi}$  is defined as  $\hat{\phi}(\chi) = \chi \circ \phi$  for  $\chi \in \hat{G}'$  and we prove that if  $\hat{\phi}$  is surjective then  $\phi$  is injective and vice-versa.

## Section 6

1. a

$R^{Reg}$ , denote the regular matrix representation of  $G$ . We have to compute the matrix  $\sum_{g \in G} R_g^{Reg}$ . We will write  $R_g$  for  $R_g^{Reg}$ , to keep things short. Suppose  $|G| = n$ . Then each  $R_g$  is a  $n \times n$  matrix. Let us fix  $1 \leq i, j \leq n$ . First we determine  $(R_g)_{ij}$ . Observe that  $(R_g)_{ij} = 1$ , if  $gg_j = g_i$ , else the entry is 0. But as there exist a unique  $g$  such that  $gg_j = g_i$ , we conclude the following:

For  $1 \leq i, j \leq n$ , there exists a unique  $G$  such that  $(R_g)_{ij} = 1$ , and for all other  $h \in G$ , the  $ij^{th}$  entry of the matrix  $R_h$  is 0. Therefore, we have that  $(\sum_{g \in G} R_g)_{ij} = 1$ . Since  $i, j$  was chose arbitrarily, we see that the matrix  $\sum_{g \in G} R_g$  is a matrix all of whose entries are 1.

## Result

2 of 2

The main idea is to calculate the  $ij$  entry of  $R_g^{Reg}$ , for each  $g \in G$ . See the solution for more details.

2. a

Let  $\rho$  be the permutation representaton of  $S_3$ , obtained from  $S_3$  acting on  $S_3$  by conjugation. Observe that  $\rho$  is a representation of dimension 6, as  $|S_3| = 6$ . Now, we have decompose  $\rho$ , into irreducible representations. Now, Let  $\chi$  denote the character of  $\rho$ . From the book, table 10.4.12, we know the irreducible characters of  $S_3$ . Let them me denoted by  $\chi_1, \chi_2, \chi_3$ , just as in the book.

Now, we first determine, the character  $\chi$ . It is enough determine on a representative of each conjugacy class. Now, if a group  $G$  acts on itself by conjugation, and  $g \in G$ , then

$$Fix(g) = \{x \in G | gxg^{-1} = x\} = C_G(g), \text{ where } C_G(g) \text{ denote the centralizer of } g \text{ in } G.$$

Now,  $\chi(g) = |Fix(g)| = |C_G(g)|$ . So  $\chi(e) = 6$ ,  $\chi((12)) = 2$ ,  $\chi((123)) = 3$ . Now, in each case we compute the inner products.

$$\langle \chi, \chi_1 \rangle = \frac{1}{6}(6 + 3 \cdot 2 + 2 \cdot 3) = 3.$$

$$\langle \chi, \chi_2 \rangle = \frac{1}{6}(6 - 3 \cdot 2 + 2 \cdot 3) = 1$$

$$\langle \chi, \chi_3 \rangle = \frac{1}{6}(6 \cdot 2 + 2 \cdot 0 + 2 \cdot 0 + 2 \cdot 0 + 3 \cdot -1 + 3 \cdot -1) = 2$$

So finally, we get  $\chi = 3 \cdot \chi_1 + \chi_2 + 2 \cdot \chi_3$ . This is the decomposition of  $\chi$ .

We refer to table 10.4.12, for the character table of  $S_3$ . We compute the character of the permutation representation of  $S_3$ , and then compute inner products with each one of the irreducible characters, to see how they occur in its decomposition.

3. a

To decompose character of tetrahedral group  $T$  on the six edges of the tetrahedron into irreducible characters;

Suppose  $\chi^e$  denote the character of the representation of the tetrahedral group  $T$ .

Consider the tetrahedron group  $T$ , it has six edges, four vertex and four faces.

Since it is known that order of  $T$  is 12 and three representations have dimension 1 and one representation has dimension 3.

So this can be written as,

$$|T| = d_1^2 + d_2^2 + d_3^2 + d_4^2$$

$$12 = 1^2 + 1^2 + 1^2 + 3^2$$

There is no normal subgroup of order 4, so index 3.

Now, the factor group is the third order cyclic group and its characters representation as shown below;

	$e$	$t$	$t'$
$S$	1	1	1
$A$	1	$\omega$	$\omega^2$
$R$	1	$\omega^2$	$\omega$

So, use this and orthonormality conditions to obtain character table of tetrahedron group:

$T$	$\zeta_1$	$3\zeta_a$	$4\zeta_b$	$4\zeta_c$
$\chi_1$	1	1	1	1
$\chi_2$	1	1	$\omega$	$\omega^2$
$\chi_3$	1	1	$\omega^2$	$\omega$
$\chi_4$	3	-1	0	0

**Thus**, there are four conjugate classes and the dimension of four irreducible representations are 1,1,1,3 and the relation between the characters of the tetrahedron group is

$$\chi^e = \chi_1 + \chi_2 + \chi_3 + 3\chi_4.$$

4. a



(a)

The objective is to identify the five conjugacy classes in the octahedral group  $O$ , and find the orders of its irreducible representations.

---

[Comment](#)

---

Step 2 of 9 ^

Consider the octahedral group which is denoted by  $O$ .

Octahedral group has order 24 and 5 conjugacy classes and 5 irreducible representations.

Degree 1 for 4 irreducible representations:

$$1 + 1 + 1 + 1 + x^2 = 24$$

This implies that,

$$x^2 = 20$$

This is not possible because 20 is not a square.

Degree 1 for 3 irreducible representations:

$$1 + 1 + 1 + x^2 + y^2 = 24$$

This implies that,

$$x^2 + y^2 = 21$$

This is not possible because 21 is not the sum of two squares.

---

[Comment](#)

---

Step 4 of 9 ^

Degree 1 for 2 irreducible representations:

$$1 + 1 + x^2 + y^2 + z^2 = 24$$

And suppose  $x \leq y \leq z$

For  $x = 3$  then  $y^2 + z^2 \geq 13$  this is impossible.

For  $x = 2$  then  $y^2 + z^2 = 18$  and  $y = 3, z = 3$  is the only possibility.

Degree 1 for 1 irreducible representations:

$$1 + x^2 + y^2 + z^2 + a^2 = 24$$

Here,  $x = 2$  then  $y^2 + z^2 + a^2 = 19$ .

This is impossible.

Hence, orders of the irreducible representation of octahedral groups are **1, 1, 2, 3, and 3** and

conjugacy classes are **1,  $2_y$ ,  $2_{xz\theta}$ ,  $3_{xxx}^+$ , and  $4_y^+$**

[Comment](#)

Step 6 of 9 ^

(b)

The objective is to decompose the provided characters into irreducible characters when group  $O$  operates the following sets:

1. Six faces of the cube.
2. Three pairs of opposite vertices.
3. Eight vertices.
4. Four pairs of opposite vertices.
5. Six pairs of opposite edges.
6. Two inscribed tetrahedral.

To decompose the characters into irreducible characters;

Decomposition of the characters into irreducible representation as shown below:

The irreducible representation of octahedral group as shown below:

Six faces of the cube:

$$\chi = \chi_1 \oplus \chi_2 \oplus \chi_3 \oplus \chi_4 \oplus 2\chi_5$$

Three pairs of opposite vertices:

$$\chi = \chi_1 \oplus \chi_2 \oplus \chi_3$$

Eight vertices:

$$\chi = \chi_1 \oplus \chi_3 \oplus 2\chi_3 \oplus 2\chi_4 \oplus 2\chi_5$$

Four pairs of opposite vertices:

$$\chi = \chi_1 \oplus \chi_2 \oplus \chi_3 \oplus \chi_4$$

Six pairs of opposite edges:

$$\chi = \chi_1 \oplus \chi_2 \oplus \chi_3 \oplus \chi_4 \oplus 2\chi_5$$

Two inscribed tetrahedral:

$$\chi = \chi_1 \oplus \chi_2$$

Hence, all these sets represents as characters into irreducible characters.

(c)

The objective is to determine the character table for  $O$ ;

Class length of  $O$  are 1,3,6,8,6 and element order are 1,2,2,3,4

Number of character represented as  $\chi_1, \chi_2, \chi_3, \chi_4, \chi_5$  and since there are 5 conjugacy classes as  $1, 2_y, 2_{x0}, 3_{xxx}, 4_y^+$ .

The character table of octahedral group as shown below:

$O$	1	1	2	3	4
	1	$2_y$	$2_{x0}$	$3_{xxx}$	$4_y^+$
$\chi_1$	1	1	1	1	1
$\chi_2$	1	1	-1	1	-1
$\chi_3$	2	2	0	-1	0
$\chi_4$	3	-1	-1	0	1
$\chi_5$	3	-1	1	0	-1

5. a

**Solution:** By the given condition the symmetric group  $S_n$  operates on  $\mathbb{C}^n$  by permuting the coordinates. We will decompose this representation explicitly into irreducible representations.

Recall that the vector  $(1, 1, \dots, 1)$ , of dimension  $n$ , is invariant under the action of  $S_n$  and spans one-dimensional trivial representation.

Let us consider the vector space

$$V := \left\{ (x_1, x_2, x_3, \dots, x_n) \mid \sum_k x_k = 0 \right\}.$$

Then it follows that by the definition of  $V$ ,  $V$  is invariant of  $S_n$ .

Now we will propose to prove that  $V$  is an irreducible representation.

If possible let us assume that it has a nontrivial invariant subspace  $X$  and

$$x := (x_1, x_2, x_3, \dots, x_n) \in X.$$

Now notice that if  $x_i = x_j$  for all  $i, j$  then we have

$$\sum x_k = 0 \implies x_k = 0 \quad \forall k.$$

Otherwise note that

$$x_i \neq x_j \quad \text{for some } i, j.$$

Therefore we have

$$\begin{aligned} & x - (i, j)x \\ &= (0, \dots, 0, x_i - x_j, 0, \dots, 0, x_j - x_i, 0, \dots, 0) \in X. \end{aligned}$$

And note that

$$\frac{1}{x_i - x_j}(x - (i, j)x) \\ = (0, \dots, 0, 1, 0, \dots, 0, -1, 0, \dots, 0) \in X.$$

Now by applying  $S_n$  to this vector, we can get all vectors of the form  $(0, \dots, 0, 1, 0, \dots, 0, -1, 0, \dots, 0)$ , which span  $V$ .

This follows that

$$X = V.$$

Hence we proved that  $V$  is an irreducible representation.

This completes the proof.

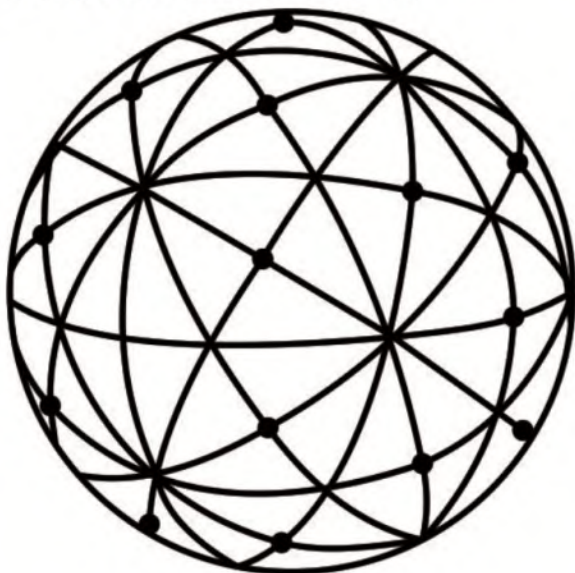
## Result

2 of 2

Considering the vector space  $V$  as the hyperplane passing through the origin and prove that  $V$  is an irreducible representation.

6. a

To decompose the characters of the representations of the icosahedral group on the sets of faces, edges, and vertices into irreducible characters;



It is known that the icosahedral group is represented by  $A_5$ .

Let  $G = A_5$

Order of  $A_5$  is 60 and  $G$  has 5 conjugacy classes with representatives as follows:

Representative	$e$	$(123)$	$(12)(34)$	$(12345)$	$(13452)$
order	1	20	15	12	12
Centralizer order	60	3	4	5	5

Here there are 5 irreducible characters. Obviously there is the trivial representation, with the trivial character  $\chi_1(a) = 1$  for all  $a \in G$ . In order to find the other characters of  $S_5$  and restrict them to  $A_5$ .

The character table of the icosahedral group  $A_5$  as shown below:

	$e$	$(123)$	$(12)(34)$	$(12345)$	$(13452)$
$\chi_1$	1	1	1	1	1
$\chi_2$	4	1	0	-1	-1
$\chi_3$	5	-1	1	0	0
$\chi_4$	3	0	-1	$\frac{1+\sqrt{5}}{2}$	$\frac{1-\sqrt{5}}{2}$
$\chi_5$	3	0	-1	$\frac{1-\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$

The icosahedron has 12 vertices and 20 faces. Five edges originate from each vertex, but each edge is on two vertices. So there are 30 edges.

## 7. a

Consider the group  $S_5$  operates by conjugation on its normal subgroup  $A_5$ .

First find the character table for  $S_5$ .

Order of  $S_5$  is 120.

Find the conjugacy classes of  $S_5$ .

$$C_1 = \{(1)\},$$

$$C_2 = \{(12)\},$$

$$C_3 = \{(1234)\},$$

$$C_4 = \{(12345)\},$$

$$C_5 = \{(12)(34)\}$$

$$C_6 = \{(12)(345)\}$$

Therefore, the complete character table for  $S_5$  is:

$S_5$	1	(12)	(123)	(1234)	(12345)	(12)(34)	(12)(345)
$\chi_1$	1	1	1	1	1	1	1
$\chi_2$	1	-1	1	-1	1	1	-1
$\chi_3$	4	2	1	0	-1	0	-1
$\chi_4$	4	-2	1	0	-1	0	1
$\chi_5$	6	0	0	0	1	-2	0
$\chi_6$	5	1	-1	-1	0	1	1
$\chi_7$	5	-1	-1	1	0	1	-1

To find character table for  $A_5$ ;

The conjugacy classes of  $A_5$  as shown below:

$$C_1 = \{(1)\},$$

$$C_2 = \{(123)\},$$

$$C_3 = \{(12)(34)\},$$

$$C_4 = \{(12345)\},$$

$$C_5 = \{(13452)\}$$

From the character table of  $S_5$  it follows that  $\chi_1|_H = \chi_2|_H = \chi_3|_H = \chi_4|_H$  and  $\chi_6|_H = \chi_7|_H$  are irreducible characters of  $A_5$ .

The character table of  $A_5$  as shown below:

$A_5$	1	(123)	(12)(34)	(12345)	(13452)
$\varphi_1$	1	1	1	1	1
$\varphi_2$	4	1	0	-1	-1
$\varphi_3$	5	-1	1	0	0
$\varphi_4$	3	0	-1	$\frac{1+\sqrt{5}}{2}$	$\frac{1-\sqrt{5}}{2}$
$\varphi_5$	3	0	-1	$\frac{1-\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$

**Thus,** It can be seen that the only irreducible character of  $S_5$  whose restriction to  $A_5$  which is not irreducible is  $\chi_3$  and  $\varphi_1, \varphi_2, \varphi_3$  are all obtained by restriction of the characters other than  $\chi_3$ .

8. a



To decompose the restriction to  $T$  of the irreducible character of  $I$  where  $T$  is tetrahedral group and  $I$  is icosahedral group;

Consider the icosahedral group  $C_5 = I$ ;

Let  $\omega$  denote a primitive fifth root of unit in  $\mathbb{C}^*$  such that  $\omega^5 = 1$ .

There are five conjugacy classes in  $C_5$  and five elements.

Moreover, to determine a representation of  $C_5$ ;

Now, to describe the image of the generator that is a fifth root of unity.

Hence, a power of  $\omega$

The character table of icosahedral group  $I$  as shown below:

$I$	0	1	2	3	4
$\gamma_1$	1	1	1	1	1
$\gamma_\omega$	1	$\omega$	$\omega^2$	$\omega^3$	$\omega^4$
$\gamma_{\omega^2}$	1	$\omega^2$	$\omega^4$	$\omega$	$\omega^3$
$\gamma_{\omega^3}$	1	$\omega^3$	$\omega$	$\omega^4$	$\omega^2$
$\gamma_{\omega^4}$	1	$\omega^4$	$\omega^3$	$\omega^2$	$\omega$

Consider the tetrahedron group  $T$ , it has six edges, four vertex and four faces.

Since it is known that order of  $T$  is 12 and three representations have dimension 1 and one representation has dimension 3.

So this can be written as,

$$|T| = d_1^2 + d_2^2 + d_3^2 + d_4^2$$

$$12 = 1^2 + 1^2 + 1^2 + 3^2$$

There is no normal subgroup of order 4, so index 3.

Now, the factor group is the third order cyclic group and its characters representation as shown below;

	$e$	$t$	$t'$
$S$	1	1	1
$A$	1	$\omega$	$\omega^2$
$R$	1	$\omega^2$	$\omega$

So, use this and orthonormality conditions to obtain character table of tetrahedron group  $T$ :

$T$	$\zeta_1$	$3\zeta_a$	$4\zeta_b$	$4\zeta_c$
$\chi_1$	1	1	1	1
$\chi_2$	1	1	$\omega$	$\omega^2$
$\chi_3$	1	1	$\omega^2$	$\omega$
$\chi_4$	3	-1	0	0

Hence, the restriction to  $I$  of the characters of  $\chi_1$  and sign equal to the trivial character, the character of  $\chi_2$  is equal to  $\gamma_\omega$ , the character of  $\chi_3$  is equal to  $\gamma_{\omega^2}$ , and the character of  $\chi_4$  is sum of  $\gamma_1, \gamma_\omega$  and  $\gamma_{\omega^2}$ .

9. a

First we prove a general result which will be of use in many problems of this book. This result actually determines the following: Given a finite group  $G$ , and its character table  $T$ , how can one find all normal subgroups of  $G$ . To do this we first prove the following claim:

**Claim 1** Suppose  $R$  is a representation of  $R : G \rightarrow GL(V)$  of degree  $d$ , and  $\chi_R$  denote the character of  $R$ . Define  $Ker(\chi_R) = \{g \in G | \chi_R(g) = d\}$ . We claim that  $Ker(\chi_R) = Ker R$ , where  $Ker R$  is the usual kernel of the homomorphism  $R$ . Indeed if  $g \in Ker R$ , then  $R_g = id_V$ . Since degree of  $R$  is  $d$ , we get  $\chi_R(g) = \text{trace}(id_V) = d$ . Therefore  $g \in Ker(\chi_R)$ . This proves  $Ker R \subseteq Ker(\chi_R)$ .

Conversely suppose  $g \in Ker(\chi_R)$ . This implies  $\chi_R(g) = d \implies |\chi_R(g)| = d$ . Now, it is a standard observation that  $\chi_R(g)$  is sum of  $d$  roots of unity. Also, it is a standard result that if  $\zeta_1, \dots, \zeta_d$  are roots of unity then  $|\zeta_1 + \dots + \zeta_d| \leq d$ , with equality iff  $\zeta_1 = \zeta_2 = \dots = \zeta_d$ . so, from this we conclude that as  $|\chi_R(g)| = d$ , then  $\chi_R = d\zeta$ , where  $\zeta$  is a root of unity. But,  $\chi_R(g) = d$  implies that  $\zeta = 1$ . Hence, we get that all the eigen values of  $R_g$  is 1. But  $R_g$  is diagonalizable (again standard observation, as minimal polynomial of  $R_g$  divides  $X^n - 1$ , where  $n$  is the order of  $g$ ), and therefore we conclude that  $R_g = id_V$ . Therefore,  $Ker(\chi_R) \subseteq Ker R$ . This completes the proof of the claim.

Suppose now that  $\chi$  is any character (not necessarily irreducible) of  $G$ . Let  $\chi = \sum_{i=1}^k n_i \chi_i$ , where  $\chi_1, \chi_2, \dots, \chi_k$  denote all the distinct irreducible characters, and  $n_i (1 \leq i \leq k)$  are non-negative integers. Let  $d_i$  denote the degrees of the character  $\chi_i$  for every  $1 \leq i \leq k$ .

**Claim 2:**  $Ker \chi = \cap \{ Ker \chi_i | n_i > 0 \}$ .

Let  $A = \cap \{ Ker \chi_i | n_i > 0 \}$ . Let  $g \in A$ . Then  $\chi_i(g) = d_i$ , for every  $i$  such that  $n_i > 0$ . If  $d$  is the degree of  $\chi$ , it is clear that  $\chi(g) = d$ . Therefore,  $g \in Ker \chi$ . This proves  $A \subseteq Ker \chi$ . Conversely, suppose  $g \in Ker \chi$ , then  $\chi(g) = d$ . So  $|\chi(g)| \leq d_i$ , claim 1 forces  $\chi_i(g) = d_i$ , whenever  $n_i > 0$ . Therefore, we have that  $g \in A$ . Thus, we have proved our second claim.

Now, we proceed towards the final claim, which tells us how to determine all normal subgroups of  $G$ , by the character table of  $G$ . Let,  $N_i = Ker \chi_i$ . It is clear the  $N_i$  are normal subgroups.

**Claim 3:** Any normal subgroup of  $G$  is a certain intersection of  $N_i$ 's.

Let  $N$  be any normal subgroup of  $G$ . Consider  $G/N$ . Let  $\rho$  denote the regular representation of  $G/N$ . Let  $\pi : G \rightarrow G/N$  be the canonical surjection, with  $\text{Ker}\pi = N$ . Consider  $\rho_G = \rho \circ \pi$ . Observe that since  $\rho$  is injective,  $\rho_G$  is a representation of  $G$ , with  $\text{Ker}\rho_G = N$ . Let  $\chi_G$  be the character of  $\rho_G$ . Then, by claim 1,  $N = \text{Ker}\chi_G$ . Now, Let

$$\chi_G = \sum_{i=1}^k n_i \chi_i$$

. Then, from claim 2, we get  $N = \text{Ker}\chi_G = \cap \{ \text{Ker}\chi_i | n_i > 0 \} = \cap \{ N_i | n_i > 0 \}$ .

So, finally it is easy to determine the  $N_i$ 's from the character table. To determine  $N_i$ , just look at the  $i^{\text{th}}$  row, and find out all elements of  $G$ , whose character entry is  $d_i$ , where  $d_i = \chi_i(1)$ . Then to determine all normal subgroups, just take all possible intersection of these  $N_i$ 's, and list the new subgroups thus obtained. This gives you all the possible normal subgroups. Now, to determine whether a group is simple, Observe that these  $N_i$ 's must be either  $G$  or  $\{e\}$ . Since the other normal subgroups are intersections of these they are also  $G$  or  $\{e\}$ . Therefore, if all the  $N_i$ 's are trivial, then the group is simple.

The following is the character table of the icosahedral group.

	0	$\pi$	$2\pi/3$	$2\pi/5$	$4\pi/5$
$\chi_1$	1	1	1	1	1
$\chi_2$	3	-1	0	$\alpha$	$\beta$
$\chi_3$	3	-1	0	$\beta$	$\alpha$
$\chi_4$	4	0	1	-1	-1
$\chi_5$	5	1	-1	0	0

Here  $\alpha = 1 + \cos(2\pi/5)$  Now,  $N_1 = \text{Ker}\chi_1 = \{g \in G | \chi_1(g) = 1\} = G$

$N_2 = \text{Ker}\chi_2 = \{g \in G | \chi_2(g) = 3\} = \{e\}$

$N_3 = \text{Ker}\chi_3 = \{g \in G | \chi_3(g) = 3\} = \{e\}$

$N_4 = \text{Ker}\chi_4 = \{g \in G | \chi_4(g) = 4\} = \{e\}$

$N_5 = \text{Ker}\chi_5 = \{g \in G | \chi_5(g) = 5\} = \{e\}$ .

Therefore, since each  $N_i (1 \leq i \leq 5)$  are trivial, we conclude that the icosahedral group  $I$ , is simple.

## Result

5 of 5

To understand how to determine from the character table, whether the group  $G$  is simple, we need to understand how to determine all normal subgroups from the character table. That is what we prove first and then as a simple application we explain how to determine whether the group  $G$  is simple. Click to see more details.

The conjugacy classes of  $A_4$  are:

The identity, 4 3-cycles represented by  $(123)$ , 4 3-cycles represented by  $(132)$ , 3 products of two disjoint transpositions, represented by  $(12)(34)$ .

[Comment](#)

#### Step 2 of 7 ^

Now,

Consider the representations of  $S_4$  and simply restrict them to  $A_4$ .

There are 4 irreducible representations, and the sum of the squares of their dimensions is 12.

$$\begin{aligned}\chi_{U|_{A_4}} &= (1, 1, 1, 1) \\ &= \chi_{U'|_{A_4}}\end{aligned}$$

This means that,  $U|_{A_4}$  and  $U'|_{A_4}$  are isomorphic.

$$\begin{aligned}\chi_{V|_{A_4}} &= (3, 0, 0, -1) \\ &= \chi_{(V \oplus U')|_{A_4}}\end{aligned}$$

This implies that,

$$(\chi_{V|_{A_4}}, \chi_{V|_{A_4}}) = 1$$

And,

So,  $V|_{A_4}$  is irreducible, with dimension 3. It is known that missing 2 representations and their dimensions must be 1, since,

$$1^2 + 3^2 + 1^2 + 1^2 = 12$$

Now, continue restricting the irreducible representations of  $S_4$  to  $A_4$ :

$$\chi_{W|_{A_4}} = (2, -1, -1, 2)$$

This implies that,

$$(\chi_{W|_{A_4}}, \chi_{W|_{A_4}}) = 2$$

This means that,  $W|_{A_4} = V_1 \oplus V_2$  with  $V_i$  irreducible representation.

[Comment](#)

#### Step 4 of 7 ^

Now find  $V_i$  by using the projection formulas.

$$\begin{aligned}(\chi_{W|_{A_4}}, \chi_{V_1|_{A_4}}) &= \frac{1}{12} (1 \cdot 2 \cdot 1 + 4 \cdot (-1) \cdot 1 + 4 \cdot (-1) \cdot 1 + 3 \cdot 2 \cdot 1) \\ &= 0\end{aligned}$$

$$\begin{aligned}(\chi_{W|_{A_4}}, \chi_{V_2|_{A_4}}) &= \frac{1}{12} (1 \cdot 2 \cdot 3 + 4 \cdot (-1) \cdot 0 + 4 \cdot (-1) \cdot 0 + 3 \cdot 2 \cdot (-1)) \\ &= 0\end{aligned}$$

This conclude that neither  $U_{|A_4}$  nor  $V_{|A_4}$  are sub representations of  $W_{|A_4}$ .

Since still have 2 representations left and at least one of them is a sub representation of  $W_{|A_4}$ .

Note that,

$$A_4 / \{1, (12)(34), (13)(24), (14)(23)\} = Z_3$$

[Comment](#)

Step 6 of 7 ^

Since  $Z_3 = \{0, 1, 2\}$  is an abelian group,

So,

It is known those irreducible representations  $\rho: Z_3 \rightarrow \mathbb{C}$  with  $\omega = e^{2\pi i/3}$ :

The trivial representation, which sends all elements into 1,

A representation which sends  $0 \rightarrow 1, 1 \rightarrow \omega, 2 \rightarrow \omega^2$ ,

A representation which sends  $0 \rightarrow 1, 1 \rightarrow \omega^2, 2 \rightarrow \omega$ ,

Let this 3 representations be  $U$ ,  $U'$  and  $U''$  respectively, use them for  $A_4$ .

It is known that,

$$\begin{aligned}\chi_U((12)(34)) &= \chi_{U'}((12)(34)) \\ &= \chi_{U''}((12)(34)) \\ &= 1\end{aligned}$$

Thus, character table as shown below:

	1	(123)	(132)	(12)(34)
$U$	1	1	1	1
$U'$	1	$\omega$	$\omega^2$	1
$U''$	1	$\omega^2$	$\omega$	1
$V$	3	0	0	-1

11. a

!!!

## Section 7

1. a



Let  $G$  be a finite group. Let  $\rho$  be a representation of  $G$ , say  $\rho : G \rightarrow GL(V)$ . It is given any  $G$ -invariant linear operator is a scalar multiple of identity. Now, Let us assume that  $\rho$  is reducible. So, by Mashke's theorem,  $\rho$  is a decomposable representation. In other words, there exists  $V \supset V_1, V_2 \neq \{0\}$ , such that  $V = V_1 \oplus V_2$ , and  $V_1, V_2$ , are both  $G$ -invariant. Now, define  $T : V \rightarrow V$ , as follows:

Suppose  $v \in V$ , then there exists unique  $v_1 \in V_1, v_2 \in V_2$ , such that  $v = v_1 + v_2$ . Then we define  $T(v) = \lambda_1 v_1 + \lambda_2 v_2$ , where  $\lambda_1 \neq \lambda_2$ , and both are non zero. Now, for  $g \in G$ , we have,

$$T(\rho_g(v)) = T(\rho_g(v_1 + v_2)) = T(\rho_g(v_1) + \rho_g(v_2)) = T(\rho_g(v_1)) + T(\rho_g(v_2)).$$

But, observe that,  $\rho_g(v_1) \in V_1, \rho_g(v_2) \in V_2$ , and therefore, from the above relation we get that,

$$T(\rho_g(v)) = \lambda_1 \rho_g(v_1) + \lambda_2 \rho_g(v_2) = \rho_g(T(v_1)) + \rho_g(T(v_2)) = \rho_g(T(v_1 + v_2)) = \rho_g(T(v))$$

Since  $v \in V$ , was chose arbitrarily,  $T \circ \rho_g = \rho_g \circ T$ , for every  $g \in G$ . So,  $T$  is  $G$ -invariant operator, but since  $\lambda_1 \neq \lambda_2$ , it is clear that  $T$  is not a scalar operator. Therefore, we have arrived at a contradiction. So,  $\rho$  is irreducible.

## Result

2 of 2

This is the converse of Schur's Lemma, and we have proved, by the method of contradiction.

## 2. a

Let  $A$  denote the standard matrix representation of  $S_3$ . Let

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Now, we have to use the operator "left multiplication by B" and the averaging process to find a  $S_3$ -invariant operator. We first show a general construction of  $G$ -invariant operator from an arbitrary linear operator, and then we will use the construction to find the desired result.

Suppose  $G$  is a finite group and  $R : G \rightarrow GL_n(\mathbb{C})$  and  $S : G \rightarrow GL_m(\mathbb{C})$  denote two matrix representation of  $G$ . Suppose  $M : \mathbb{C}^n \rightarrow \mathbb{C}^m$  be a linear operator and let  $M$  (using the same notation) denote its matrix with respect to the standard basis. In other words,  $M$  is given by "left multiplication by  $M$ ". Clearly  $M$  is a  $m \times n$  matrix. Now, we construct  $\tilde{M}$ , which is  $G$ -invariant matrix. Consequently the linear operator, given by "left multiplication by  $\tilde{M}$ " will be the  $G$ -invariant operator. Define the following:

$$\tilde{M} = \frac{1}{|G|} \sum_{g \in G} S_{g^{-1}} M R_g$$

Now, Let  $h \in G$ . Then,

$$\begin{aligned} S_{h^{-1}} \left( \sum_{g \in G} S_{g^{-1}} M R_g \right) R_h &= \sum_{g \in G} S_{h^{-1}g^{-1}} M R_{gh} = \\ \sum_{g \in G} S_{(hg)^{-1}} M R_{hg} &= \sum_{g \in G} S_{g^{-1}} M R_g. \end{aligned}$$

The last line is true because left multiplication by an element of the group, actually permutes all the elements of the group. This proves that  $\tilde{M}$  is  $G$ -invariant.



Now, we come to our scenario. In our case  $G = S_3, R = S = A$ . First we explicitly write down  $A: A: S_3 \rightarrow GL_2(\mathbb{C})$ . Let  $c = \cos(2\pi/3), s = \sin(2\pi/3)$ .

$$A_{(123)} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}, A_{(12)} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We also get,

$$A_{(132)} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, A_{(13)} = \begin{bmatrix} c & s \\ s & -c \end{bmatrix}, A_{(23)} = \begin{bmatrix} c & -s \\ -s & -c \end{bmatrix}$$

Now, we find  $\tilde{A} = \frac{1}{|G|} \sum_{g \in G} A_{g^{-1}} B A_g$ . So,  $A_e B A_e = B$ .

$$A_{(12)} B A_{(12)} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

$$A_{(13)} B A_{(13)} = \begin{bmatrix} c^2 + sc & cs - c^2 \\ sc + s^2 & s^2 - cs \end{bmatrix}$$

$$A_{(23)} B A_{(23)} = \begin{bmatrix} c^2 - cs & -c^2 - cs \\ s^2 - sc & s^2 + sc \end{bmatrix}$$

$$A_{(132)} B A_{(132)} = \begin{bmatrix} c^2 + cs & c^2 - cs \\ -s^2 - sc & s^2 - sc \end{bmatrix}$$

$$A_{(123)} B A_{(123)} = \begin{bmatrix} c^2 - sc & c^2 + cs \\ sc - s^2 & s^2 + sc \end{bmatrix}$$

Therefore,

$$\tilde{A} = \frac{1}{6} \begin{bmatrix} 4c^2 + 2 & 0 \\ 0 & 4s^2 + 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{2}{3} \end{bmatrix}$$

Therefore,  $\tilde{A}$  is the required  $S_3$ — invariant matrix. In other words, the required  $S_3$ — invariant operator is given by

**"left multiplication by  $\tilde{A}$ "**

## Result

3 of 3

We have first proved a general way to find a  $G$ — invariant operator using any arbitrary operator. Then we have used that general result to find the linear operator in the case of the standard representation of  $S_3$ , and the arbitrary linear operator given by the matrix  $B$ . See the solution for more details.

3. a

Let  $R$  denote the matrix representation of  $S_3$ , given by

$$R_{(123)} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, R_{(12)} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

Let,

$$B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider the operator  $\phi : \mathbb{C} \rightarrow \mathbb{C}^3$ , given by "left multiplication by  $B$ ".

Now, we have to use the operator  $\phi$  and the averaging process to find a  $S_3$ -invariant operator. We first show a general construction of  $G$ -invariant operator from an arbitrary linear operator, and then we will use the construction to find the desired result.

Suppose  $G$  is a finite group and  $S : G \rightarrow GL_n(\mathbb{C})$  and  $R : G \rightarrow GL_m(\mathbb{C})$  denote two matrix representation of  $G$ . Suppose  $M : \mathbb{C}^n \rightarrow \mathbb{C}^m$  be a linear operator and let  $\tilde{M}$  (using the same notation) denote its matrix with respect to the standard basis. In other words,  $\tilde{M}$  is given by "left multiplication by  $M^*$ ". Clearly  $\tilde{M}$  is a  $m \times n$  matrix. Now, we construct  $\tilde{M}$ , which is  $G$ -invariant matrix. Consequently the linear operator, given by "left multiplication by  $\tilde{M}^*$ " will be the  $G$ -invariant operator. Define the following:

$$\tilde{M} = \frac{1}{|G|} \sum_{g \in G} R_{g^{-1}} M S_g$$

Now, Let  $h \in G$ . Then,

$$R_h^{-1} \tilde{M} S_h = \frac{1}{|G|} R_{h^{-1}} \left( \sum_{g \in G} R_{g^{-1}} M S_g \right) R_h = \frac{1}{|G|} \sum_{g \in G} R_{h^{-1}g^{-1}} M S_{gh} = \frac{1}{|G|} \sum_{g \in G} R_{(hg)^{-1}} M S_{hg} = \frac{1}{|G|} \sum_{g \in G} R_{g^{-1}} M S_g = \tilde{M}. \text{ The last line is true because left multiplication by an element of the group, actually permutes all the elements of the group. This proves that } \tilde{M} \text{ is } G\text{-invariant.}$$

Now, with this general construction in mind, in our case we have,

$G = S_3, R : S_3 \rightarrow GL_3(\mathbb{C})$ , as described at the beginning of the problem. Also,  $S : S_3 \rightarrow GL_1(\mathbb{C})$ , be the sign map, that is,

$$S(e) = 1, S((12)) = S((23)) = S((13)) = -1, S((123)) = S((132)) = 1.$$

Also, we have  $\tilde{M} = B$ , where  $B$  is described in the beginning. Now, we calculate  $\tilde{M}$ , according to the above general construction. Observe that  $\tilde{M}$  will be  $3 \times 1$  matrix. We first  $R$  explicitly. We already have

$$R_{(123)} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, R_{(12)} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

Also, we have the following

$$R_{(13)} = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, R_{(23)} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, R_{(132)} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Now, we can use our general theory. So,

$$R_e B S_e = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, R_{(12)} B S_{(12)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, R_{(23)} B S_{(23)} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$R_{(13)} B S_{(13)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, R_{(132)} B S_{(132)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, R_{(123)} B S_{(123)} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$\tilde{M} = \frac{1}{6} \sum_{g \in G} R_{g^{-1}} B S_g$ . Therefore,

$$\tilde{M} = \frac{1}{6} \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

Therefore the required  $G$ -invariant operator constructed  $\phi$ , say  $\psi : \mathbb{C} \rightarrow \mathbb{C}^3$ , is given by

**left multiplication by  $\tilde{M}$**

## Result

3 of 3

The  $G$ -invariant linear operator  $\psi : \mathbb{C} \rightarrow \mathbb{C}^3$ , constructed from  $\phi$  is given by

**left multiplication by  $\tilde{M}$**

, where,

$$\tilde{M} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

The solution is exactly same as the previous problem. See the solution for more details.

4. a

Let  $G$  be a finite group and  $\rho : G \rightarrow GL(V)$ , be a representation. Let  $C$  be a conjugacy class. Define the following operator:

$$T : V \rightarrow V, \text{ as } T = \sum_{g \in C} \rho_g.$$

We have to prove that  $T$  is  $G$ -invariant. Let  $h \in G$ . Observe that  $\rho_h \rho_g \rho_h^{-1} = \rho_h \rho_g \rho_{h^{-1}} = \rho_{hgh^{-1}}$ . Now, if  $g \in C$ , then  $hgh^{-1} \in C$ . So,

$$\rho_h \circ T \circ \rho_h^{-1} = \rho_h \circ \left( \sum_{g \in C} \rho_g \right) \circ \rho_h^{-1} = \sum_{g \in C} \rho_{hgh^{-1}}.$$

Therefore as  $hgh^{-1} \in C$ , and the fact that  $hg_1h^{-1} = hg_2h^{-1}$ , implies  $g_1 = g_2$ , says that the above sum  $\sum_{g \in C} \rho_{hgh^{-1}}$ , actually runs over all the elements of the conjugacy class  $C$ . Therefore, we conclude that  $\sum_{g \in C} \rho_{hgh^{-1}} = \sum_{g \in C} \rho_g$ . Hence from the above calculation we have that

$$\rho_h \circ T \circ \rho_h^{-1} = \sum_{g \in C} \rho_g = T.$$

Hence,  $\rho_h T = T \rho_h$ , for every  $h \in G$ . Therefore,  $T$  is  $G$ -invariant.

## Result

2 of 2

We have just used the definition of  $G$ -invariant operator, to obtain a proof. See the solution for more details

5. a

Let  $G$  be a finite group and  $\rho : G \rightarrow GL(V)$ , be a representation. Let  $\chi$  be a character of  $G$  not necessarily the character of  $\rho$ . Define the following operator:

$$T : V \rightarrow V, \text{ as } T = \sum_{g \in G} \chi(g) \rho_g.$$

We have to prove that  $T$  is  $G$ -invariant. Let  $h \in G$ .

$$\rho_h \circ T \circ \rho_h^{-1} = \rho_h \circ \left( \sum_{g \in G} \chi(g) \rho_g \right) \circ \rho_h^{-1} = \sum_{g \in G} \chi(g) \rho_{hgh^{-1}} \quad (1)$$

We observe here that as  $\chi(g)$ , is a scalar, it comes out in the above calculation.

Let  $C_1, C_2, \dots, C_k$ , be the conjugacy classes of  $G$ , and let  $g_1, g_2, \dots, g_k$ , be their respective representatives. Now, we first rewrite the operator  $T$  as follows:

$$T = \sum_{g \in G} \chi(g) \rho_g = \sum_{i=1}^k \chi(g_i) \left( \sum_{g \in C_i} \rho_g \right) \quad (2)$$

This rewriting must be clear because of the fact that  $\chi$  is a class function, and therefore on each conjugacy class it is constant. What we have done is that we have clubbed all the terms with the same coefficient.

Now, from equation (1), observe that the term  $\chi(g)$  is attached to the operator  $\rho_{hgh^{-1}}$ . This show that  $\chi(g_i)$  is attached to  $\rho_{hgh^{-1}}$ . But, as  $g$  runs over  $C_i$ ,  $hgh^{-1}$ , also runs over  $C_i$ , with constant coefficient  $\chi(g_i)$ , for each  $1 \leq i \leq k$ . Then we have that eqn (1), can be rewritten as

$$\rho_h T \rho_h^{-1} = \sum_{g \in G} \chi(g) \rho_{hgh^{-1}} = \sum_{i=1}^k \chi(g_i) \left( \sum_{g \in C_i} \rho_g \right) \quad (3)$$

Therefore, from eqn (2), and eqn (3), we get  $T \rho_h = \rho_h T$ , for every  $h \in G$ , and therefore  $T$  is  $G$ -invariant.

## Result

2 of 2

To prove this one uses the definition of  $G$ -invariant, and the fact that a character of  $G$  is a class function. See the proof more details.

## 6. a

Let  $\mathcal{M}$  denote the linear space of  $m \times n$  matrices.  $\dim(\mathcal{M}) = mn$ . Let  $A, B$  are matrices of orders  $m \times m$ , and  $n \times n$  respectively. Define the linear operator  $F : \mathcal{M} \rightarrow \mathcal{M}$ , defined by  $F(M) = AMB$ . In lemma 10.8.1, it has been proved that  $\text{trace}(F) = \text{trace}(A)\text{trace}(B)$ . We will reprove this lemma, now by explicitly calculating a matrix of  $F$ . For that we fix a basis  $\mathcal{B}$  of  $\mathcal{M}$  which consists of the matrices  $(E_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ , where the matrix  $E_{ij}$  has the standard definition that  $(i, j)$  position is 1, others have entry 0. Let  $A = (a_{ij}), B = (b_{kl})$ . Now, we will find the matrix of  $F$ , with respect to  $\mathcal{B}$ . For fix  $i, j$ ,  $F(E_{ij}) = AE_{ij}B$ .

Observe that  $X = AE_{ij}$ , can be described as follows, the  $j^{\text{th}}$  column of  $X$ , same as the  $i^{\text{th}}$  column of  $A$ . All the other columns are zero columns. Then  $(XB)_{rs} = (a_{ri}b_{js})$ . Therefore if  $F_{\mathcal{B}}$  denote the matrix of  $F$ , then  $(F_{\mathcal{B}})_{ij} = a_{i(d+1)}b_{tj}$ , where  $j = nd + t$ . Here, we assume that if  $t = 0$ , we actually take  $t = n$  and  $d = d - 1$ . Now, observe that  $\text{trace}(F) = \sum_{i=1}^{mn} (F_{\mathcal{B}})_{ii} = \sum_{i=1}^{mn} a_{(d+1)(d+1)}b_{tt}$ , where  $i = nd + t$ , and if  $t = 0$ , we replace  $t = n$ , and  $d = d - 1$ . Observe that,  $\text{trace}(F_{\mathcal{B}}) = (\sum_{i=1}^m a_{ii})(\sum_{j=1}^n b_{jj}) = \text{trace}(A)\text{trace}(B)$

## Result

2 of 2

This is an alternative proof of Lemma 10.8.1. We write the matrix  $F$ , by suitable choosing an order basis  $\mathcal{B}$ , which has been defined in the answer. See the solution for more details

## Section 8

1. a

Consider the following locus:

$$x_0^2 + \dots + x_3^2 \leq r^2$$

To calculate the four-dimensional volume of the 4-ball  $\mathbf{B}^4$  of radius  $r$  in  $\mathbb{R}^4$ ;

Use  $x_0 = r \cos \theta, x_1 = r \sin \theta$  and  $x_2 = 0, x_3 = 0$  to parameterize the slices of the 4-ball  $\mathbf{B}^4$ .

Integration by slicing provided a recursive formula for the volume of the ball:

$$\text{vol}_n(r) = 2 \int_0^{\pi/2} \text{vol}_{n-1}(r \cos \theta) r \cos \theta d\theta$$

The zero balls consist of one point.

So,

$$\text{vol}_0(r) = 1$$

For  $n = 1$ ,

$$\begin{aligned} \text{vol}_1(r) &= 2 \int_0^{\pi/2} \text{vol}_0(r \cos \theta) r \cos \theta d\theta \\ &= 2 \int_0^{\pi/2} 1 \cdot r \cos \theta d\theta \\ &= 2r (\sin \theta) \Big|_0^{\pi/2} \\ &= 2r \end{aligned}$$

For  $n = 2$ ,

$$\begin{aligned} \text{vol}_2(r) &= 2 \int_0^{\pi/2} \text{vol}_1(r \cos \theta) r \cos \theta d\theta \\ &= 2 \int_0^{\pi/2} 2r \cos \theta r \cos \theta d\theta \\ &= 4r^2 \int_0^{\pi/2} \cos^2 \theta d\theta \\ &= 4r^2 \int_0^{\pi/2} \left( \frac{\cos 2\theta + 1}{2} \right) d\theta \\ &= 2r^2 \left( \frac{\sin 2\theta}{2} + \theta \right) \Big|_0^{\pi/2} \\ &= 2r^2 \times \frac{\pi}{2} \\ &= \pi r^2 \end{aligned}$$



For  $n = 3$ ;

$$\begin{aligned}
 \text{vol}_3(r) &= 2 \int_0^{\pi/2} \text{vol}_2(r \cos \theta) r \cos \theta d\theta \\
 &= 2 \int_0^{\pi/2} \pi r^2 \cos^2 \theta r \cos \theta d\theta \\
 &= 2\pi r^3 \int_0^{\pi/2} \cos^3 \theta d\theta \\
 &= 2\pi r^3 \left( \sin \theta - \frac{1}{3} \sin^3 \theta \right) \Big|_0^{\pi/2} \\
 &= 2\pi r^3 \times \frac{2}{3} \\
 &= \frac{4}{3} \pi r^3
 \end{aligned}$$

For  $n = 4$ ;

$$\begin{aligned}
 \text{vol}_4(r) &= 2 \int_0^{\pi/2} \text{vol}_3(r \cos \theta) r \cos \theta d\theta \\
 &= 2 \int_0^{\pi/2} \frac{4}{3} \pi r^3 \cos^3 \theta r \cos \theta d\theta \\
 &= \frac{8}{3} \pi r^4 \int_0^{\pi/2} \cos^4 \theta d\theta \\
 &= \frac{8}{3} \pi r^4 \left( \frac{3}{8} \theta + \left( \frac{1}{4} \right) \sin 2\theta + \frac{1}{32} \sin 4\theta \right) \Big|_0^{\pi/2} \\
 &= \frac{8}{3} \pi r^4 \left( \frac{3}{8} \times \frac{\pi}{2} \right) \\
 &= \frac{1}{2} \pi^2 r^4
 \end{aligned}$$

Therefore, four-dimensional volume of the 4-ball  $\mathbf{B}^4$  of radius  $r$  in  $\mathbb{R}^4$  is:

$$\boxed{\frac{1}{2} \pi^2 r^4}$$

Now, to check answer by the differentiating with respect to  $r$ , to get

Suppose that volume of the 4-ball  $\mathbf{B}^4$  of radius  $r$  in  $\mathbb{R}^4$  denoted by  $\text{vol}_4(r)$  then,

$$\text{vol}_4(r) = \frac{1}{2} \pi^2 r^4$$

Differentiate with respect to  $r$ ,

$$\begin{aligned}
 \frac{\partial \text{vol}_4(r)}{\partial r} &= \frac{4}{2} \pi^2 r^3 \\
 &= 2\pi^2 r^3
 \end{aligned}$$

Thus, for a 4-dimensional ball the answer is that the boundary of three dimensional balls has an area of  $2\pi^2 r^3$ .

**Hence**, answer is correct.

2. a



**Operation 10.9.3** as shown below;

$$[Pf](u, v) = f(ua - v\bar{b}, ub + v\bar{a})$$

Where,

$$P = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

[Comment](#)

Step 2 of 4 ^

To verify the associative law  $[Q[Pf]] = [(QP)f]$  for the operation 10.9.3;

Suppose that matrix  $P$  and  $Q$  as shown below,

$$P = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

And,

$$Q = \begin{bmatrix} c & d \\ -\bar{d} & \bar{c} \end{bmatrix}$$

Suppose that the elements of the matrices are in complex number.

Now apply operation in  $[Q[Pf]]$ :

$$\begin{aligned} [Q[Pf]](u, v) &= [Q]f(ua - v\bar{b}, ub + v\bar{a}) \\ &= \begin{bmatrix} c & d \\ -\bar{d} & \bar{c} \end{bmatrix} f(ua - v\bar{b}, ub + v\bar{a}) \\ &= f(c(ua - v\bar{b}) - \bar{d}(ub + v\bar{a}), d(ua - v\bar{b}) + \bar{c}(ub + v\bar{a})) \\ &= f((ca - \bar{d}b)u - (c\bar{b} + \bar{a}\bar{d})v, (ad + b\bar{c})u + (-\bar{b}d + \bar{a}\bar{c})v) \end{aligned}$$

Since elements in the complex number, so

$$[Q[Pf]](u, v) = f((ca + db)u - (-cb + ad)v, (ad - bc)u + (bd + ac)v)$$

Determine  $QP$ :

$$\begin{aligned} QP &= \begin{bmatrix} c & d \\ -\bar{d} & \bar{c} \end{bmatrix} \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix} \\ &= \begin{bmatrix} ac - d\bar{b} & bc + \bar{a}d \\ -\bar{d}a - \bar{b}\bar{c} & -b\bar{d} + \bar{a}\bar{c} \end{bmatrix} \end{aligned}$$

Again apply operation in  $[(QP)f]$ :

$$\begin{aligned} [(QP)f](u, v) &= \begin{bmatrix} ac - d\bar{b} & bc + \bar{a}d \\ -\bar{d}a - \bar{b}\bar{c} & -b\bar{d} + \bar{a}\bar{c} \end{bmatrix} f(u, v) \\ &= f((ac - d\bar{b})u - (\bar{d}a + \bar{b}\bar{c})v, (bc + \bar{a}d)u + (-b\bar{d} + \bar{a}\bar{c})v) \end{aligned}$$

Since elements in the complex number so,

$$\begin{aligned} [(QP)f](u, v) &= f((ac + db)u - (da - bc)v, (-bc + ad)u + (bd + ac)v) \\ &= f((ca + db)u - (-cb + ad)v, (ad - bc)u + (bd + ac)v) \end{aligned}$$

**Hence**, verified the associative property  $[Q[Pf]] = [(QP)f]$ .

3. a

Theorem:

Define a surjective homomorphism  $\rho: SU_2 \rightarrow SO_3$ , the spin homomorphism. Its kernel is the center  $\{\pm 1\}$  of  $SU_2$ .

[Comment](#)

## Step 2 of 3 ^

Show that the orthogonal representation  $SU_2 \rightarrow SO_3$  is irreducible.

Define an orthogonal representation  $\rho: SU_2 \rightarrow SO_3$ .

Suppose that  $f \in SU_2$  on the lie algebra by the representation of adjoint.

Suppose that adjoint representation denoted by  $\gamma$ .

Then,

$$\gamma(f)v = fvf^{-1}$$

It is known that inner product on  $SU_2$  is:

$$\langle v, w \rangle = -\text{tr}(vw)$$

Because, inner product  $\langle f v, f w \rangle = \langle v, w \rangle$ .

$$\begin{aligned} \langle f v, f w \rangle &= -\text{tr}(f v f^{-1} f w f^{-1}) \\ &= -\text{tr}(v w) \\ &= \langle v, w \rangle \end{aligned}$$

So, the orthogonal representation  $\rho: SU_2 \rightarrow SO_3$  is holomorphic.

It is easy to see that  $\pm 1$  are in kernel of  $\rho$  in  $SU_2$  and also it is commute with all  $2 \times 2$  matrix in  $SO_3$  and only scalar multiples of the identity operator can commute with every matrix in  $SU_2$ .

So, the kernel of  $\rho$  is exactly  $\{\pm 1\}$ .

Thus,  $\rho$  is irreducible.

**Hence**, the orthogonal representation  $SU_2 \rightarrow SO_3$  is irreducible.

## 4. a

To decompose the associated complex representation into irreducible representation;

Left multiplication defines a representation of  $SU_2$  on the space  $\mathbb{R}^4$  with coordinates  $x_0, \dots, x_3$ .

Defines a bijective correspondence of  $SU_2$  with the unit 3-sphere  $\{x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1\}$  in  $\mathbb{R}^4$  as shown below,

$$SU_2 \rightarrow \mathbf{S}^3$$

This is defined by,

$$P = \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix} \rightarrow (x_0, x_1, x_2, x_3)$$

There is exactly one irreducible representation of  $SU_2$  because number of linearly independent in matrix  $P$  is one.

Define a representation  $\rho: SU_2 \rightarrow S^3$  by,

$$\rho(P) = (x_0, x_1, x_2, x_3)$$

To show that this representation is irreducible, so prove that this is closed under addition and multiplication.

Define a map  $\phi: V \rightarrow V$  commuted with all  $\rho(P)$  that is,

$$\phi(\rho(P)v) = \rho(P)\phi(v)$$

For all  $P \in SU_2$  and  $v \in V$ .

Then,  $\phi = \lambda \times \text{identity}_v$  for some  $\lambda \in \mathbb{C}$  and  $\rho(P) = \lambda \times \text{identity}_v$ .

**Hence**, the representation is irreducible.

5. a

**Theorem 10.9.14:**

The characters of  $SU_2$  that are orthonormal  $\langle \chi_m, \chi_n \rangle = 0$  if  $m \neq n$ , and  $\langle \chi_m, \chi_n \rangle = 1$

$$\langle \chi_m, \chi_n \rangle = \frac{1}{2\pi^2} \int_0^\pi \chi_m(\theta) \chi_n(\theta) \text{vol}_2(\sin \theta) d\theta$$

[Comment](#)

Step 2 of 4 ^

To determine the irreducible representations of the rotation group  $SO_3$ ;

Consider the irreducible representation of the group  $G$  of rotations in  $\mathbb{R}^3$ . These are orthogonal transformations of determinant 1.

The character of such a representation as shown below:

$$\chi_m(\theta) = \sum_{j=-m}^m \exp[i j \theta]$$

Where 1 and  $e^{\pm i\theta}$  are the eigenvalues.

The rotation  $H$  around  $x$ -axis is calculated as,

$$H = z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z}$$

The polynomial are harmonic in two and therefore three variables. This representation has dimension  $(2m+1)$  or less. In fact its dimension is only  $(2m+1)$ .

Apply theorem 10.9.14:

$$\begin{aligned} \langle \chi_m, \chi_n \rangle &= \frac{1}{2\pi^2} \int_0^\pi \chi_m(\theta) \chi_n(\theta) \text{vol}_2(\sin \theta) d\theta \\ &= \frac{1}{2\pi^2} \int_0^\pi \left( \frac{\alpha^{m+1} - \alpha^{-(m+1)}}{\alpha - \alpha^{-1}} \right) \left( \frac{\alpha^{n+1} - \alpha^{-(n+1)}}{\alpha - \alpha^{-1}} \right) (-\pi(\alpha - \alpha^{-1})^2) d\theta \\ &= -\frac{1}{2\pi} \int_0^\pi (\alpha^{m+n+2} - \alpha^{-(m+n+2)}) d\theta + \frac{1}{2\pi} \int_0^\pi (\alpha^{m-n} - \alpha^{n-m}) d\theta \\ &= 1 \end{aligned}$$

This representation forms the orthogonality of the sine functions over  $(0, \pi)$ , the characters  $\chi_n$  are orthonormal and the representation of rotation group  $SO_3$  is irreducible.

6. a

Suppose  $G$  is the circle group  $\{e^{i\theta}\}$  and assume that all representation to be differentiable functions of  $\theta$ .

[Comment](#)

Step 2 of 10 ^

(a)

To show that there exists a positive definite  $G$  invariant Hermitian form on  $V$ .

Let a positive definite, symmetric bilinear form  $\{ , \}$  on  $V$ .

To find such a form it is enough to fix an isomorphism of  $V$  with  $\mathbb{C}^n$  and consider the standard positive definite symmetric bilinear form on  $\mathbb{C}^n$ .

Now define an averaging form:

$$\langle v, w \rangle = \frac{1}{|G|} \sum_{g \in G} \{ \rho(g)v, \rho(g)w \}$$

To check the above form is symmetric, positive definite and  $G$ -invariant.

That is, symmetric follows from the fact that  $\{ , \}$  is symmetric:

$$\begin{aligned} \langle v, w \rangle &= \frac{1}{|G|} \sum_{g \in G} \{ \rho(g)v, \rho(g)w \} \\ &= \frac{1}{|G|} \sum_{g \in G} \{ \rho(g)w, \rho(g)v \} \end{aligned}$$

Hence, the form of averaging is symmetric:

Now to check the form of averaging is positive:

To show that  $\langle v, v \rangle$  is positive so,

$$\begin{aligned} \langle v, v \rangle &= \frac{1}{|G|} \sum_{g \in G} \{ \rho(g)v, \rho(g)v \} \\ &> 0 \end{aligned}$$

Hence, the form of averaging is positive.

[Comment](#)

Step 4 of 10 ^

To check the form of averaging is  $G$ -invariant:

Since the expression is a sum of positive numbers, rearranging the summation to verify the  $G$  invariant, for any element  $h \in G$ , the right multiplication by  $h$  gives a permutation of  $G$ :

$$\begin{aligned} \langle \rho(h)v, \rho(h)w \rangle &= \frac{1}{|G|} \sum_{g \in G} \{ \rho(g)\rho(h)v, \rho(g)\rho(h)w \} \\ &= \frac{1}{|G|} \sum_{g \in G} \{ \rho(gh)v, \rho(gh)w \} \\ &= \langle v, w \rangle \end{aligned}$$

Hence, the form of averaging is  $G$ -invariant.

(b)

**Maschke's theorem:**

Every representation of a finite group  $G$  on a nonzero, finite dimensional complex vector space is a direct sum of irreducible representations.

[Comment](#)

---

Step 6 of 10 ^

To show that Maschke's theorem for  $G$ ;

Suppose that  $\rho$  is irreducible, then there is nothing to prove.

Suppose that  $\rho$  is not irreducible and let  $W < V$  a  $\rho$ -invariant subspace.

Claim:  $W^\perp$  is orthogonal vector with respect to the form  $\langle \cdot, \cdot \rangle$ .

In part (a), the representation of arranged form is  $\rho$ -invariant.

To check that  $z \in W^\perp$  and  $g \in G$ , then  $\rho(g)z \in W^\perp$  or for every  $w \in W$ ,

$$\langle \rho(g)z, w \rangle = 0$$

This true when,

$$\langle \rho(g)z, w \rangle = \langle \rho(g^{-1})\rho(g)z, \rho(g^{-1})w \rangle$$

Because  $\langle \cdot, \cdot \rangle$  is  $\rho$ -invariant and  $\rho$  is a representation

$$\langle \rho(g)z, w \rangle = \langle z, \rho(g^{-1})w \rangle$$

Since  $\rho(g^{-1})w \in W$  because  $W$  is  $\rho$ -invariant subspace and  $z \in W^\perp$ , so

$$\langle \rho(g)z, w \rangle = 0$$

Since  $V$  is finite dimensional complex vector space, so

This implies that  $\rho$  splits as a direct sum of two representations of irreducible.

**Hence**, finite dimensional complex vector space is a direct sum of irreducible representations.

(c)

To describe the representations of  $G$  in terms of one-parameter groups;

Suppose that  $\rho_n$  be irreducible representation of  $G$  such that,

$$\rho_n(e^{i\theta}) = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$$

Where,  $n \in \mathbb{Z}^+$

To prove that the irreducible representations are one-dimensional;

Let  $\rho_n$  be an irreducible complex representation of  $G$ .

Since  $G$  is abelian, it is known that

$$\begin{aligned}\rho_n(g)\rho_n(h)v &= \rho_n(gh)v \\ &= \rho_n(hg)v \\ &= \rho_n(h)\rho_n(g)v\end{aligned}$$

For all  $v \in V$

Apply Schur's Lemma,

$$\rho_n(g)v = av$$

For any  $g \in G$ , where  $a$  is some complex scalar.

Therefore, all subspace of  $V$  is  $G$ -invariant,

Irreducibility of  $V$  implies that the  $G$ -invariant are the trivial spaces  $\{0\}$  and  $V$  itself. Since  $\dim V > 1$  requires  $V$  having a non-trivial subspace, but this contradicts the statement.

Therefore,  $V$  is 1-dimensional.

**Hence**, the irreducible representations are one-dimensional.

(d)

To verify the orthogonality relation, using an analogue of the Hermitian product orthogonal.

The Hermitian product as shown below:

$$\langle \chi_m, \chi_n \rangle = \frac{1}{|G|} \int_G \overline{\chi_m(g)} \chi_n(g) dv$$

Since it is known that,

$$\begin{aligned}\langle \chi_m, \chi_n \rangle &= \frac{1}{2\pi^2} \int_0^\pi (\chi_m(\theta) \chi_n(\theta) + \text{vol}_2(\sin \theta)) d\theta \\ &= \frac{1}{2\pi^2} \int_0^\pi \left( \frac{\alpha^{m+1} - \alpha^{-(m+1)}}{\alpha - \alpha^{-1}} \right) \left( \frac{\alpha^{n+1} - \alpha^{-(n+1)}}{\alpha - \alpha^{-1}} \right) (-\pi(\alpha - \alpha^{-1})^2) d\theta \\ &= -\frac{1}{2\pi} \int_0^\pi (\alpha^{m+n+2} + \alpha^{-(m+n+2)}) d\theta + \frac{1}{2\pi} \int_0^\pi (\alpha^{m-n} + \alpha^{n-m}) d\theta\end{aligned}$$

If  $m \neq n$  then,

$$\langle \chi_m, \chi_n \rangle = 0$$

**Hence**, the orthogonality relation verified.

7. a

To determine the irreducible representations of the orthogonal group  $O_2$ ;

[Comment](#)

Step 2 of 4 ^

Since it is known that the group  $SO_2$  is abelian.

So it can be written as the group  $O_2$  is a semi-direct product of two abelian groups.

$$O_2 = SO_2 \times \left\{ \begin{pmatrix} \pm 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

Now identify  $SO_2$  with  $\mathbb{C} = \{e^{i\theta} : \theta \in \mathbb{R}\}$ ;

$$A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

This implies that,

$$A_\theta \rightarrow e^{i\theta}$$



Now identify the dual of the Abelian group  $SO_2$  with the group  $\mathbb{Z}$  of integers  $n \in \mathbb{Z}$  this tends to

$$\varepsilon_n.$$

Where,

$$\varepsilon_n : e^{i\theta} \mapsto e^{in\theta}$$

Consider the matrix  $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ ; it is conjugative on group  $SO_2$  which is  $SO_2 \cong \mathbb{C}$  and taking

inverse, that is  $e^{i\theta} \mapsto e^{-i\theta}$ .

Hence, dual map  $n$  from  $-n$ , that is  $n \mapsto -n$ .

Now define orbit of  $\varepsilon_n$  and stabilizer of  $\varepsilon_n$  by,

$$\text{Orbit of } \varepsilon_n = \begin{cases} \{\varepsilon_n, \varepsilon_{-n}\} & n \neq 0 \\ \{1\} & n = 0 \end{cases}$$

And,

$$\text{Stabilizer of } \varepsilon_n = \begin{cases} \{1\} & n \neq 0 \\ \left\{ \begin{pmatrix} \pm 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} & n = 0 \end{cases}$$

For  $n > 0$  the representation  $\rho_n$  is irreducible, and for  $n = 0$  that  $\rho_n$  splits into two one-dimensional representations, that is,  $\det O_2 = \pm 1$  and the trivial representation.  $\rho_n = 1 \oplus \det O_2$

Hence,  $\rho_n \cong \rho_{-n}$  and  $\rho_n = 1 \oplus \det O_2$ , these are isomorphic to all irreducible representations of the orthogonal group  $O_2$ .

## Miscellaneous Problem

1. a

!!!

2. a

Suppose  $G$  is a finite group, with exactly three irreducible representations of degree 1, 2, 3 resp. Now, by main theorem 10.4.6, we have that  $|G| = 1^2 + 2^2 + 3^2 = 14$ . But also by the main theorem, we know that degree of a irreducible character must divide the order of the group. Now, observe that  $G$  has a irreducible representation of degree 3, but  $3 \nmid |G| = 14$ . Therefore, we conclude that such a group  $G$  doesn't exist.

### Result

2 of 2

We have shown that a finite group doesn't exists with the given hypothesis. The main ingredient is to use the main theorem 10.4.6.

3. a

Let  $\rho$  be a representation of  $G$ . So,  $\rho : G \rightarrow GL(V)$ .

(a) Suppose  $x \in G$ , and define  $\rho' : G \rightarrow GL(V)$ , as  $\rho'(g) = \rho(xgx^{-1})$ . We claim that  $\rho'$  is a representation. This is easy to see. Let  $g_1, g_2 \in G$ . Then  $\rho'(g_1g_2) = \rho(xg_1g_2x^{-1}) = \rho(xg_1x^{-1}xg_2x^{-1}) = \rho(xg_1x^{-1})\rho(xg_2x^{-1}) = \rho'(g_1)\rho'(g_2)$ . Therefore, we have proved that  $\rho'$  is a representation of  $G$ . Next, we claim that  $\rho$  and  $\rho'$  are isomorphic. For this purpose, Let  $\chi, \chi'$ , be the characters of  $\rho, \rho'$  respectively. Suppose  $g \in G$ . Now  $\chi'(g) = \text{trace}(\rho(xgx^{-1})) = \text{trace}(\rho(x)\rho(g)\rho(x)^{-1}) = \text{trace}(\rho(g)) = \chi(g)$ . So, we have  $\chi = \chi'$ , and hence the two representations are isomorphic.

## Step 2

2 of 4

(b) Suppose  $\phi$  is a automorphism of  $G$ . Define  $\rho'(g) = \rho(\phi(g))$ . We first claim that  $\rho'$  is a representation of  $G$ . Let  $g_1, g_2 \in G$ . Now  $\rho'(g_1g_2) = \rho(\phi(g_1g_2)) = \rho(\phi(g_1)\phi(g_2)) = \rho(\phi(g_1))\rho(\phi(g_2)) = \rho'(g_1)\rho'(g_2)$ . This proves that  $\rho'$  is a representation of  $G$ . Now, consider  $\rho : \mathbb{Z}_6 \rightarrow \mathbb{C}^*$ , given by  $\bar{1} \mapsto e^{i\pi/6}$ . Here  $\mathbb{Z}_6$  is the residue class of integers modulo 6, which is a cyclic group, under the operation of usual addition. Now, let  $\phi : \mathbb{Z}_6 \rightarrow \mathbb{Z}_6$  denote the automorphism  $\bar{a} \mapsto \bar{5a}$ . Then, observe that  $\rho' : \mathbb{Z}_6 \rightarrow \mathbb{C}^*$ , is given by  $\bar{1} \mapsto e^{5\pi i/6}$ . It is clear that  $\rho \neq \rho'$ . Since  $\rho, \rho'$  are one-dimensional representation, we conclude that  $\rho$  and  $\rho'$  are not isomorphic.

(c) Let  $\sigma : G \rightarrow \mathbb{C}^*$  is a one-dimensional representation. Define  $\rho' : G \rightarrow GL(V)$ , as  $\rho'(g) = \sigma(g)\rho(g)$ . We claim again that  $\rho'$  is a representation. Let  $g_1, g_2 \in G$ . Now  $\rho'(g_1g_2) = \sigma(g_1g_2)\rho(g_1g_2) = \sigma(g_1)\sigma(g_2)\rho(g_1)\rho(g_2) = \sigma(g_1)\rho(g_1)\sigma(g_2)\rho(g_2) = \rho'(g_1)\rho'(g_2)$ . Thus our claim is proved. Now, we show that  $\rho$  and  $\rho'$  need not be isomorphic. Let  $\chi$  and  $\chi'$ , be the characters of the representation of  $\rho$ , and  $\rho'$  respectively. Observe that  $\chi'(g) = \sigma(g)\chi(g)$ . Now, take  $G = S_4$ . Suppose  $\rho$  is the standard representation of  $S_4$ . Let the representatives of the conjugacy classes of  $S_4$  be given by  $e, (12), (123), (1234), (12)(34)$ . Then,  $\chi(e) = 3, \chi((12)) = 1, \chi((123)) = 0, \chi((1234)) = -1, \chi((12)(34)) = -1$ . Let  $\sigma$  denote the sign map of  $S_4$ . Then, observe that  $\chi'((1234)) = 1$ . Therefore, in this case  $\chi \neq \chi'$ , and hence  $\rho$  and  $\rho'$  are non-isomorphic.

## Result

4 of 4

In (a), we have proved that  $\rho$  and  $\rho'$  are isomorphic. In the other two cases, we have provided counterexample to show that  $\rho$  and  $\rho'$  are non-isomorphic. See the solution for more details.

## 4. a

Let  $G$  be a finite group and  $\rho : G \rightarrow GL(V)$ , be an irreducible representation. Suppose  $z \in Z(G)$ . Then,  $zg = gz$ , for every  $g \in G$ . Consider  $\rho_z : V \rightarrow V$ , the image of  $z$  under  $\rho$ . We claim that  $\rho_z$  is  $G$ -invariant. Indeed, for  $g \in G, \rho_g\rho_z = \rho_{gz} = \rho_z\rho_g$ . Now, applying Schur's lemma, as  $\rho$ , is irreducible, we conclude that  $\rho_z = \lambda I$ , for some  $\lambda \in \mathbb{C}^*$ .

Conversely, assume that  $z \in G$ , and  $\rho_z$  is scalar multiple of identity for every irreducible representation  $\rho$  of  $G$ . We have to show that  $z \in Z(G)$ . Now, first observe that if  $g \in G$ , then  $\rho_z\rho_g = \rho_g\rho_z$ , as  $\rho_z$  is scalar. This implies that  $\rho_{zg} = \rho_{gz}$ , and hence  $\rho_{zgz^{-1}g^{-1}} = I$ . If  $\chi$  is the character of  $\rho$  and  $\rho$  has degree  $d$ , then  $\chi(zgz^{-1}g^{-1}) = d$ .

Suppose  $\rho_1, \rho_2, \dots, \rho_k$  be the irreducible representation of  $G$ , with degree  $d_1, d_2, \dots, d_k$ , and characters  $\chi_1, \chi_2, \dots, \chi_k$ . From the discussion, we have  $\chi_i(zgz^{-1}g^{-1}) = d_i$ , for every  $1 \leq i \leq k$ . Now, let  $\chi_{reg}$ , be the character of the regular representation of  $G$ . Then we can say,

$$\chi_{reg}(t) = \sum_{i=1}^k d_i \chi_i(t) \quad \dots (1)$$

Now, taking  $t = zgz^{-1}g^{-1}$ , in equation (1), we see that

$$\chi_{reg}(zgz^{-1}g^{-1}) = \sum_{i=1}^k d_i^2 = |G| \quad \dots (2)$$

So, from the above equation, we see that  $\chi_{\text{reg}}(zgz^{-1}g^{-1}) = |G|$ . But we know that

$$\chi_{\text{reg}}(g) = \begin{cases} |G| & g = 1 \\ 0 & , \text{otherwise} \end{cases}$$

This implies that  $zgz^{-1}g^{-1} = 1$ , implies that  $zg = gz$ . Since,  $g$  was chosen arbitrarily, we conclude that  $z \in Z(G)$ . This completes the proof.

## Result

2 of 2

For the first part, we show that  $\rho_z$  is a  $G$ -invariant operator, and use Schur's lemma, and for the second, we use the regular representation of  $G$ .

5. a

**Given:**  $A$  and  $B$  are two matrices such that some positive power of each matrix is the identity and satisfying

$$AB = BA.$$

**To Prove:** There exists an invertible matrix  $P$  such that both of  $PAP^{-1}$  and  $PBP^{-1}$  are diagonal matrices.

**Proof:** First we show that both of  $A$  and  $B$  are diagonalizable.

By the given condition there exists a natural number  $n$  such that

$$A^n = I, \quad \text{where } I \text{ is the identity matrix.}$$

We know that a linear operator is diagonalizable if and only if its minimal polynomial splits into distinct linear factors.

Since each entries of  $A$  are in  $\mathbb{C}$  and the minimal polynomial is given by

$$f(x) = x^n - 1$$

which splits in the field of complex numbers, so by the above theory  $A$  is diagonalizable.

By the similar thought of argument  $B$  is also diagonalizable.

**Lemma:** Let  $A$  and  $B$  be two linear operators of a vector space  $V$  such that  $AB = BA$  and  $A, B$  are diagonalizable, then  $A$  and  $B$  are simultaneously diagonalizable.

**Proof of the Lemma:** Since  $A, B$  are both diagonalizable,  $V$  is the direct sum of the eigenspaces for  $A$  and  $B$ . Let  $\lambda$  be an eigenvalue for  $B$  and  $E_\lambda$  be the  $\lambda$ -eigenspace of  $B$  in  $V$ . Since  $A$  commutes with  $B$  we have

$$A(E_\lambda) \subseteq E_\lambda.$$

Therefore each  $A$  and  $B$  restricts to a linear operator on the subspace  $E_\lambda$  and the linear operators  $A|_{E_\lambda}$  and  $B|_{E_\lambda}$  commutes.

Now notice that  $A|_{E_\lambda}$  is diagonalizable on  $E_\lambda$ . Then there exists a basis vector for  $E_\lambda$  consisting of the eigenvectors for  $A|_{E_\lambda}$ .

Now note that the elements of this basis for  $E_\lambda$  are eigenvectors for  $B|_{E_\lambda}$  also, since all non-zero vectors of  $E_\lambda$  are eigenvectors for  $B$ .

Therefore  $B|_{E_\lambda}$  and  $A|_{E_\lambda}$  are diagonalizable.

Recall that the vector space  $V$  is the direct sum of the eigenspaces  $E_\lambda$  of  $A$  and  $B$ , so stringing together simultaneous eigenbases of  $B|_{E_\lambda}$  and  $A|_{E_\lambda}$  as  $\lambda$  runs over the eigenvalues of  $B$  gives a simultaneous eigenbasis of  $V$  for  $A$  and  $B$ .

This completes the proof of the Lemma.

Now from the lemma it is obvious that  $A$  and  $B$  are simultaneously diagonalizable, since  $AB = BA$  and  $A, B$  are both diagonalizable.

In other words, there exists an invertible matrix  $P$  such that both of  $PAP^{-1}$  and  $PBP^{-1}$  are diagonal matrices. This completes the proof.

## Result

3 of 3

Considering  $A$  and  $B$  as linear operators of a vector space  $V$  we have shown that  $A$  and  $B$  are simultaneously diagonalizable.

## 6. a

Let  $\rho$  be a irreducible representation of  $G$ . Suppose  $\phi_1, \phi_2 : V \times V \rightarrow \mathbb{C}$ , be two  $G$ -invariant positive definite Hermitian form on  $V$ . Now let  $V^*$ , denote the dual space of  $V$ , that is the set of all linear functionals on  $V$ . Define the following map

$\psi : V \rightarrow V^*$ , defined by for any  $v \in V$ ,  $\phi(v) : V \rightarrow \mathbb{C}$ , is given by  $\phi(v)(w) = \phi_1(w, v)$ . Observe that  $\psi$  is a bijective linear map from  $V$  to  $V^*$ . The map is bijection because the Hermitian form is positive definite, and therefore necessarily non-degenerate. One can define another map  $\theta : V \rightarrow V^*$ , exactly as before, but know using  $\phi_2$ , instead of  $\phi_1$ . Again, by the same reason  $\theta$  is a bijective linear map. Now, with the above setup consider,

$\theta^{-1} \circ \psi : V \rightarrow V$ . Suppose  $v \in V$ . Then,  $\psi(v) : V \rightarrow V^*$ , is given by  $\psi(v)(w) = \phi_1(w, v)$ . Now,  $\psi(v) \in V^*$ . Suppose  $w \in V$ , be such that  $\theta(w) = \psi(v)$ . Then, we have  $\theta^{-1}(\psi(v)) = w$ . So, we get, given  $v \in V$ ,  $\theta^{-1} \circ \psi(v) = w$ . Observe that  $\theta(w)(x) = \phi_2(x, w) = \psi(v)(x) = \phi_1(x, v)$ . So, from the relation  $\theta(w) = \psi(v)$ , we obtain that for any  $x \in V$ ,  $\phi_2(x, w) = \phi_1(x, v)$ .



Now, with these information we have collected, we claim that  $\delta = \theta^{-1} \circ \psi$  is  $G$ -invariant operator on  $V$ . For  $g \in G$ , consider the linear map  $\rho_g : V \rightarrow V$ . Then,  $\delta \circ \rho_g(v) = \delta(\rho_g(v)) = w$  (say). Then, by above, we have that for each  $x \in V$ ,  $\phi_2(x, w) = \phi_1(x, \rho_g(v))$ . Now, observe that  $\phi_2(\rho_g(\rho_{g^{-1}}(x)), \rho_g(\rho_{g^{-1}}(w))) = \phi_1(\rho_g(\rho_{g^{-1}}(x)), \rho_g(v))$ . Since,  $\phi_1, \phi_2$ , are  $G$ -invariant, we get from the above equality that  $\phi_2(\rho_{g^{-1}}(x), \rho_{g^{-1}}(w)) = \phi_2(\rho_{g^{-1}}(x), v)$ . Since,  $\rho_{g^{-1}}$ , is invertible, we can easily deduce that for every  $y \in V$ , we have the equality,  $\phi_2(y, \rho_{g^{-1}}(w)) = \phi_1(y, v)$ . This again yields that  $\delta(v) = \rho_{g^{-1}}(w)$ . Now, we get  $\rho_g \circ \delta(v) = \rho_g(\delta(v)) = \rho_g(\rho_{g^{-1}}(w)) = w$ . So, finally we have the equality,

$$\delta \circ \rho_g = \rho_g \circ \delta$$

, for every  $g \in G$ . So, we have that  $\delta = \theta^{-1} \circ \psi$  is  $G$ -invariant operator. Now, since  $\rho$  is irreducible, by Schur's Lemma, we have that any  $G$ -invariant operator on  $V$ , is a scalar multiple of identity. So, we conclude that there exists  $\lambda \in \mathbb{C}$ , such that  $\theta^{-1} \circ \psi = \lambda I$ . So,  $\psi = \lambda \theta$ . This shows, for  $v, w \in V$ , we have  $\phi_1(v, w) = \lambda \phi_2(v, w)$ . Hence, we have proved that the  $G$ -invariant Hermitian form is unique upto scalars.

## Result

3 of 3

The idea is to use Schur's lemma. We start with two  $G$ -invariant Hermitian forms  $\phi_1$  and  $\phi_2$ , and we prove that  $\phi_2 = \lambda \phi_1$ , for some scalar  $\lambda$ . So, the final answer is that the  $G$ -invariant Hermitian form is unique upto scalars. See the solution for more details.

## 7. a

**Solution:** We will now explain the commutator subgroup of a group  $G$  in terms of the character table. The commutator subgroup of a group  $G$  is defined by

$$[G] := \{x^{-1}y^{-1}xy \mid x, y \in G\}.$$

We claim that the rows in the character table of  $G$  with 1 in the first column are precisely the characters lifted from  $[G]$ , a normal subgroup of  $G$ .

Let us assume  $\chi$  is a character of  $G$ , is lifted from  $\chi'$  a character of  $G/[G]$  then we have

$$\chi = \chi' \circ \pi.$$

This shows that  $\chi$  is one dimensional and  $\chi(1) = 1$ .

Conversely, if we assume that  $\chi(1) = 1$  then we have to find a map

$$\chi' : G/[G] \rightarrow \mathbb{C} - \{0\}$$

by the assignment

$$\chi = \chi' \circ \pi.$$

This provides us

$$\chi'(x[G]) = \chi(x).$$

Since our definition of  $\chi'$  involves a choice of coset representative, we have to show that this is a well-defined map.

So we will propose to show that the image is independent of the coset representative we choose

$$\chi(x) = \chi(xg) \quad \forall g \in G.$$

This follows that

$$\chi(g) = 1 \quad \forall g \in G.$$

So basically we have  $[G]$  is a subgroup of  $\text{Ker}(\chi')$ .

This is obvious since  $G/\text{Ker}(\chi')$  is a subgroup of  $\mathbb{C} - \{0\}$  and so is abelian and it yields that  $[G]$  is a subgroup of  $\text{Ker}(\chi')$ .

Now observe that the characters lifted from the irreducible character of  $G/[G]$  and all distinct and so the character table of  $G$  contains exactly order of  $G/[G]$  rows of dimension one.

Now we claim that

$$[G] = \cap_{\chi(x)=1} \text{ker}(\chi).$$

So we have proved that  $[G]$  is a subgroup of  $\text{ker}(\chi)$  for all one-dimensional character  $\chi$ .

Now we will prove the other part.

Now note that all irreducible characters of  $G/[G]$  are of the form  $\chi'$  where

$$\chi' \circ \pi = \chi$$

is a one-dimensional character of  $G$ .

Now if  $x \in G$  satisfies  $\chi(x) = 1$  for every irreducible character of  $G$  then  $x = 1$ . Therefore we have

$$\begin{aligned} g \in \cap_{\chi(1)=1} \text{ker}(\chi) &\implies \chi(g) = 1, \quad \text{for all one-dimensional character } \chi \\ &\implies \chi'(g[G]) = \chi' \circ \pi(g) = \chi(g) = 1 \\ &\implies \chi([G]) = [G] \\ &\implies g \in [G]. \end{aligned}$$

This follows that

$$\cap_{\chi(x)=1} \text{ker}(\chi) \subseteq [G].$$

### Step 3

3 of 4

Consequently we have

$$\cap_{\chi(x)=1} \text{ker}(\chi) = [G].$$

This completes the solution.

### Result

4 of 4

Considering the commutator group  $[G]$  we have prove that  $\cap_{\chi(x)=1} \text{ker}(\chi) = [G]$  where  $\chi$  is a character of  $G$ .

8. a



**Solution:** Let  $G$  be a finite, non-abelian simple group not of prime order.

We will propose to prove that  $G$  has no nontrivial representation of dimension 2.

Let us consider a non-trivial representation

$$\rho : G \rightarrow GL_2(\mathbb{C}).$$

Now notice that by the given condition since  $G$  is simple and the representation  $\rho$  is nontrivial, we must have

$$\ker \rho = \ker \chi = (e),$$

where  $\chi$  represents the character of  $\rho$ .

Now recall the Feit-Thompson Theorem.

The Feit-Thompson Theorem states that **every finite group of odd order is solvable**.

Thus by the Feit-Thompson Theorem  $|G|$  is even.

Therefore

$$|G| = 2n, \quad \text{where } n \in \mathbb{Z}.$$

So by Cauchy's Theorem,  $G$  must have an element  $x$  of order 2.

Now let us define a map

$$\bar{\rho} : G \rightarrow \mathbb{C} - \{0\}$$

by the assignment

$$\bar{\rho}(g) = \det(\rho(g)).$$

It is clear to observe that  $\bar{\rho}$  is a homomorphism, hence it gives a degree 1 representation of  $G$ .

And basically degree 1 representation is a trivial one.

It follows that

$$\det(\rho(g)) = 1 \quad \forall g \in G.$$

Since  $\rho$  is a representation of dimension 2 we have

$$\rho(x)^2 = \text{Id}.$$

Therefore the possible set of eigenvalues of  $\rho(x)$  are  $\{1, 1\}$ ,  $\{1, -1\}$ , and  $\{-1, -1\}$ .

Now  $\{1, 1\}$  can not be set of eigenvalues of  $\rho(x)$  since  $\ker \chi = (e)$ .

Now we know that determinant is equals to the product of all its eigenvalue.

If  $\{1, -1\}$  is the set of eigenvalues of  $\rho(x)$  then we must have

$$\det(\rho(x)) = -1.$$

Which is an impossibility. Therefore  $\{1, -1\}$  can not be set of eigenvalues of  $\rho(x)$ .

Thus, the eigenvalues of  $\rho(x)$  are  $\{-1, -1\}$ .

Therefore the characteristic polynomial of  $\rho(x)$  is given by  $(X + 1)^2$ , and  $\rho(x)$  also satisfies  $X^2 - 1$ . Since the minimal polynomial of  $\rho(x)$  must divide both of these, it follows  $\rho(x)$  satisfies  $X + 1$ .

It follows that

$$\rho(x) = -\text{identity}.$$

Now  $\rho(x)$  commutes with any matrix, since  $\rho(x)$  is a scalar multiple of the identity.

Therefore for any  $g \in G$ , we have

$$\begin{aligned} \rho(g)\rho(x) &= \rho(x)\rho(g) \\ \implies \rho(gxg^{-1}x^{-1}) &= \text{identity}. \end{aligned}$$

Therefore kernel of  $\rho$  is trivial.

This yield's that

$$\begin{aligned} gxg^{-1}x^{-1} &= e \quad \text{for all } g \in G \\ \implies x &\in Z(G). \end{aligned}$$

This shows that  $Z(G)$  is non-trivial.

But recall that  $Z(G)$ , the center of the group  $G$ , is always a normal subgroup of  $G$ .

But by the given condition  $G$  is simple, and since  $Z(G)$  is non-trivial, we must have

$$G = Z(G).$$

Which contradicts the fact that  $G$  is non-abelian.

This proves that  $\rho$  is trivial.

This completes the proof.

---

## Result

3 of 3

Considering a non-trivial representation  $\rho$  of  $G$  we contradicts the fact that  $G$  is non-abelian, which follows the result.

9. a

Let  $G$  be a finite group and  $H$  be a subgroup of index 2 in  $G$ . Suppose  $a \notin H$ . Then  $H$  and  $aH$  are two distinct cosets of  $H$  in  $G$ . Also,  $H$  is necessarily normal in  $G$  (This is a standard result in group theory, which is very easy to prove, by using the definition). Now,  $S : H \rightarrow GL_n$  be a matrix representation of  $H$ . The induced representation  $indS : G \rightarrow GL_{2n}$  is defined as

$$(indS)_h = \begin{bmatrix} S_h & 0 \\ 0 & S_{a^{-1}ha} \end{bmatrix}, (indS)_g = \begin{bmatrix} 0 & S_{ga} \\ S_{a^{-1}g} & 0 \end{bmatrix}$$

where,  $h \in H, g \in aH$ . Observe that as  $H$  is normal  $a^{-1}ha \in H$ . First we prove that  $indS$  is indeed a matrix representation of  $G$ . To do that we will consider cases. Suppose  $x, y \in G$ .

**Case 1:** Suppose  $x, y \in H$ . In this case  $xy \in H$

$$(indS)_{xy} = \begin{bmatrix} S_{xy} & 0 \\ 0 & S_{a^{-1}(xy)a} \end{bmatrix}$$

Now,

$$\begin{aligned} (indS)_x \cdot (indS)_y &= \begin{bmatrix} S_x & 0 \\ 0 & S_{a^{-1}xa} \end{bmatrix} \begin{bmatrix} S_y & 0 \\ 0 & S_{a^{-1}ya} \end{bmatrix} = \begin{bmatrix} S_x S_y & 0 \\ 0 & S_{a^{-1}xa} S_{a^{-1}ya} \end{bmatrix} \\ &= \begin{bmatrix} S_{xy} & 0 \\ 0 & S_{a^{-1}(xy)a} \end{bmatrix} = (indS)_{xy} \end{aligned}$$

We have,  $S_x S_y = S_{xy}$ , since  $S$  is a matrix representation.

**Case 2:** Suppose  $x, y \in aH$ . Let  $x = ah_1, y = ah_2$ . Therefore,  $xy = ah_1ah_2 = a^2h_3h_2$ , where  $h_1a = ah_3$ , for some  $h_3 \in H$ . This is because  $aH = Ha = G \setminus H$ . Since  $G/H$  is a group of order 2, we have  $aH^2 = H \implies a^2 \in H$ . Therefore, we finally get  $xy = a^2h_3h_2 \in H$ . Also, observe that  $xa, ya \in H$ .

$$(indS)_{xy} = \begin{bmatrix} S_{xy} & 0 \\ 0 & S_{a^{-1}(xy)a} \end{bmatrix}$$

Now,

$$\begin{aligned} (indS)_x \cdot (indS)_y &= \begin{bmatrix} 0 & S_{xa} \\ S_{a^{-1}x} & 0 \end{bmatrix} \begin{bmatrix} 0 & S_{ya} \\ S_{a^{-1}y} & 0 \end{bmatrix} = \begin{bmatrix} S_{xa}S_{a^{-1}y} & 0 \\ 0 & S_{a^{-1}x}S_{ya} \end{bmatrix} \\ &= \begin{bmatrix} S_{xaa^{-1}y} & 0 \\ 0 & S_{a^{-1}(xy)a} \end{bmatrix} = \begin{bmatrix} S_{xy} & 0 \\ 0 & S_{a^{-1}(xy)a} \end{bmatrix} = (indS)_{xy} \end{aligned}$$

Since,  $xa, a^{-1}y \in H$ , we have  $S_{xa}S_{a^{-1}y} = S_{xaa^{-1}y} = S_{xy}$ . Similarly, as  $ya, a^{-1}x \in H$ , we have  $S_{a^{-1}x}S_{ya} = S_{a^{-1}(xy)a}$ .

**Case 3:** Let  $x \in H, y \in aH$ . Let  $y = ah_1$ . Therefore,  $xy = xah_1 = ahh_1 \in aH$ . Since  $aH = Ha$ , we can write  $xa = ah$ , for some  $h \in H$ . Also, observe that  $ya \in H$ , as  $ya = ah_1a = aah_2 = a^2h_2 \in H$ . We have  $h_1a = ah_2$ , for some  $h_2 \in H$ , as  $aH = Ha$ .

$$(indS)_{xy} = \begin{bmatrix} 0 & S_{xya} \\ S_{a^{-1}xy} & 0 \end{bmatrix}$$

Now,

$$\begin{aligned} (indS)_x \cdot (indS)_y &= \begin{bmatrix} S_x & 0 \\ 0 & S_{a^{-1}xa} \end{bmatrix} \begin{bmatrix} 0 & S_{ya} \\ S_{a^{-1}y} & 0 \end{bmatrix} = \begin{bmatrix} 0 & S_x S_{ya} \\ S_{a^{-1}xa} S_{a^{-1}y} & 0 \end{bmatrix} \\ &= \begin{bmatrix} S_{xaa^{-1}y} & 0 \\ 0 & S_{a^{-1}(xy)a} \end{bmatrix} = \begin{bmatrix} 0 & S_{xya} \\ S_{a^{-1}xy} & 0 \end{bmatrix} = (indS)_{xy} \end{aligned}$$

This is because as  $ya \in H, x \in H, S_x S_{ya} = S_{xya}$ . Also, as  $a^{-1}xa, a^{-1}y \in H$ , we have,  $S_{a^{-1}xa} S_{a^{-1}y} = S_{a^{-1}xaa^{-1}y} = S_{a^{-1}xy}$ . so, finally, considering all the cases together we conclude that  $(indS)_x \cdot (indS)_y = (indS)_{xy}$ , for every  $xy \in G$ , and hence  $indS$  is a matrix representation of  $G$ .

(b) Next we calculate the character of  $\text{ind}S$ . Let  $\chi_S$  be the character of  $S$  on  $H$ . Let  $\chi_{\text{ind}S}$  be the character of  $\text{ind}S$ . Suppose  $h \in H$ . Then,  $\chi_{\text{ind}S}(h) = \text{trace}(S_h) + \text{trace}(S_{a^{-1}ha}) = \chi_S(h) + \chi_S(a^{-1}ha)$ . Let  $g \in aH$ , then  $\chi_{\text{ind}S}(g) = 0$ . So, we get

$$\chi_{\text{ind}S}(g) = \begin{cases} \chi_S(g) + \chi_S(a^{-1}ga) & g \in H \\ 0 & g \notin H \end{cases}$$

Now, suppose  $R$  is a representation of  $G$ ,  $R : G \rightarrow GL_n$ . We can restrict  $R$  to  $H$ , which is a representation of  $H$ . We call it  $\text{res}R$ . Observe that the character of  $\text{res}R$  is same as  $R$ , evaluated on  $H$ . Now,  $\text{ind}S : G \rightarrow GL_{2n}$ , has already been described. We restrict this representation to  $H$ , we call it  $\text{res}(\text{ind}S)$ . Let  $\chi_{\text{res}(\text{ind}S)}$  denote the character of  $\text{res}(\text{ind}S)$ . Denote by  $S' : H \rightarrow GL_n$ , the conjugate representation of  $S$ , defined by  $S'_h = S_{a^{-1}ha}$ . Let  $\chi_{S'}$  denote the character of  $S'$ . Then, for  $h \in H$ ,  $\chi_{S'}(h) = \chi_S(a^{-1}ha)$ .

Observe that  $\text{res}(\text{ind}S)(h) = (\text{ind}S)_h$ , and

$$(\text{ind}S)_h = \begin{bmatrix} S_h & 0 \\ 0 & S_{a^{-1}ha} \end{bmatrix}$$

So, it is clear that  $\text{res}(\text{ind}S) = S \oplus S'$ .

(c) We start with a representation  $R : G \rightarrow GL_n$  of  $G$ . Let  $\chi_R$  denote its character. Now we have to prove Frobenius reciprocity, that is  $\langle \chi_{\text{ind}S}, \chi_R \rangle = \langle \chi_S, \chi_{\text{res}R} \rangle$ . Now,

$$\langle \chi_{\text{ind}S}, \chi_R \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_{\text{ind}S}(g) \overline{\chi_R(g)} = \frac{1}{|G|} \sum_{h \in H} (\chi_S(h) + \chi_S(a^{-1}ha)) \overline{\chi_R(h)}.$$

Now, Let  $a^{-1}ha = h_1$ . Then  $h = ah_1a^{-1}$ . Therefore, continuing from the above equation, we can write

$$\begin{aligned} \langle \chi_{\text{ind}S}, \chi_R \rangle &= \frac{1}{|G|} \sum_{h \in H} \chi_S(h) \overline{\chi_R(h)} + \frac{1}{|G|} \sum_{h \in H} \chi_S(h_1) \overline{\chi_R(ah_1a^{-1})} \\ &= \frac{1}{|G|} \sum_{h \in H} \chi_S(h) \overline{\chi_R(h)} + \frac{1}{|G|} \sum_{h \in H} \chi_S(h_1) \overline{\chi_R(h_1)} = \frac{1}{|G|} \sum_{h \in H} \chi_S(h) \overline{\chi_R(h)} + \frac{1}{|G|} \sum_{h \in H} \chi_S(h) \overline{\chi_R(h)} \\ &= \frac{2}{|G|} \sum_{h \in H} \chi_S(h) \overline{\chi_R(h)} = \frac{1}{|H|} \sum_{h \in H} \chi_S(h) \overline{\chi_R(h)} = \langle \chi_S, \chi_{\text{res}R} \rangle \end{aligned}$$

This is because as  $R$  is a representation of  $G$ ,  $\chi_R$  is a class function in  $G$ , and hence even  $\chi_R(ah_1a^{-1}) = \chi_R(h_1)$ . Also, we can replace  $h_1$  by  $h$  once more, because,  $h_1 = a^{-1}ha$ , and hence as  $h$  varies over  $H$ ,  $h_1$ , also varies over  $H$ .

Hence we have proved  $\langle \chi_{\text{ind}S}, \chi_R \rangle = \langle \chi_S, \chi_{\text{res}R} \rangle$ .

(d) Now, suppose we assume that  $S$  is irreducible representation of  $H$ . Therefore, we have  $\langle \chi_S, \chi_S \rangle = 1$ . We first assume that  $S$ , and its conjugate representation  $S'$  are not isomorphic. We are required to prove that  $\text{ind}S$  is irreducible representation on  $G$ .

**Claim:**  $S'$  is irreducible. To, prove this we will use that  $S$  is irreducible. We have, for  $h \in H$ ,  $\chi_{S'}(h) = \chi_S(a^{-1}ha)$ . Now, we compute  $\langle \chi_{S'}, \chi_{S'} \rangle$ .

$$\langle \chi_{S'}, \chi_{S'} \rangle = \frac{1}{|H|} \sum_{h \in H} \chi_S(a^{-1}ha) \overline{\chi_S(a^{-1}ha)} = \frac{1}{|H|} \sum_{h \in H} \chi_S(h) \overline{\chi_S(h)} = \langle \chi_S, \chi_S \rangle = 1$$

The reason for going from the second expression to the third expression is that as  $h$  varies over distinct elements of  $H$ ,  $a^{-1}ha$  also varies over distinct elements of  $H$ .

Now, we have proved our claim.



Now we prove that  $\text{ind}S$  is irreducible. For that,

$$\langle \chi_{\text{ind}S}, \chi_{\text{ind}S} \rangle = \langle \chi_S, \chi_{\text{res}(\text{ind}S)} \rangle = \langle \chi_S, \chi_S + \chi_{S'} \rangle$$

The last equality is because of the fact that in (b), we have proved that  $\text{res}(\text{ind}S) = S \oplus S'$ , and therefore,  $\chi_{\text{res}(\text{ind}S)} = \chi_S + \chi_{S'}$ . So, continuing from the last equation, we have

$$\langle \chi_{\text{ind}S}, \chi_{\text{ind}S} \rangle = \langle \chi_S, \chi_S \rangle + \langle \chi_S, \chi_{S'} \rangle = 1$$

This is because, we know that  $\langle \chi_S, \chi_S \rangle = 1$ , and since  $S'$  is irreducible and we have assumed that  $S$  is not isomorphic to  $S'$ , we have  $\langle \chi_S, \chi_{S'} \rangle = 0$ .

Therefore, we have proved that  $\text{ind}S$  is irreducible representation of  $G$ .

Now, we assume that  $S$  and  $S'$  are isomorphic. Therefore,  $\chi_S = \chi_{S'}$ . So,

$$\langle \chi_{\text{ind}S}, \chi_{\text{ind}S} \rangle = \langle \chi_S, \chi_S \rangle + \langle \chi_S, \chi_{S'} \rangle = 2$$

. If we write  $\text{ind}S$  as sum of the irreducible representation of  $G$ , say

$$\chi_{\text{ind}S} = \sum_{i=1}^k n_i \chi_i$$

where,  $\chi_i$  are distinct irreducible characters of  $G$ , and  $n_i$ , their multiplicities occurring in  $\text{ind}S$ , for  $1 \leq i \leq k$ . Observe  $n_i$  are nonnegative integers. Then

$$\langle \chi_{\text{ind}S}, \chi_{\text{ind}S} \rangle = \sum_{i=1}^k n_i^2 = 2$$

. This shows that there exists  $1 \leq i \neq j \leq k$ , such that  $n_i = n_j = 1$ , and the rest are all 0. This further shows that,

$$\chi_{\text{ind}S} = \chi_i + \chi_j$$

. Hence, we have proved that  $\text{ind}S$  is the sum of two non-isomorphic representations of  $G$ .

## Result

7 of 7

This problem is about the induced representation. Given a representation  $S : H \rightarrow GL_n$ , and  $[G : H] = 2$ , one can construct a representation  $\text{ind}S$  of  $G$ . This is a standard way of constructing a representation on a subgroup to a representation of the bigger group. This problem explores the behaviour of induced representation. Click to see detailed solution.

10. a

Let  $G$  be a finite group, and  $H$  a subgroup of index 2. We have already proved the properties of induced representation of a representation of  $H$ . We will refer to the previous problem when needed.

(a) Observe that  $G = H \cup aH$ , where  $a \notin H$ . Let  $R$  be a matrix representation of  $G$ . Define another representation  $R'$ , as follows:

$$R'(g) = \begin{cases} R_g & \text{if } g \in H \\ -R_g & , \text{otherwise} \end{cases}$$

Let  $\chi_R, \chi_{R'}$  be the character of  $R$  and  $R'$ . Now,

$$\chi_{R'}(g) = \begin{cases} \chi_R(g) & \text{if } g \in H \\ -\chi_R(g) & \text{otherwise} \end{cases}$$

Suppose  $R, R'$  are isomorphic. Then  $\chi_R = \chi_{R'}$ . This implies  $\chi_R(g) = 0$ , for  $g \in aH$ . Hence the character of  $R$  is zero on  $aH$ , which is not equal to  $H$ . Conversely suppose that,  $\chi_R(g) = 0$  for every  $g \in aH$ . Then it is clear from the definition of  $\chi_{R'}$  above that,  $\chi_R = \chi_{R'}$ . This proves that  $R, R'$  are isomorphic. Hence (a) is proved.

Let us now prove (b). Recall from the previous problem, that if  $R : G \rightarrow GL_n$  is a representation and  $S : H \rightarrow GL_m$  is another representation, then Frobenius reciprocity say that  $\langle \chi_{ind S}, \chi_R \rangle = \langle \chi_S, \chi_{res R} \rangle$ . The notations are same as in previous problem.

Now, we have to prove that  $ind(res R) = R \oplus R'$ . Now,

$$\langle \chi_{ind(res R)}, \chi_R \rangle = \langle \chi_{res R}, \chi_{res R} \rangle$$

. This shows that  $R$  is a constituent of  $ind(res R)$ . Again, by using the previous problem, one can write down  $\chi_{ind(res R)}$  explicitly. We will do that now. Observe  $\chi_{res R}(h) = \chi_R(h)$ , for every  $h \in H$ . Then by using previous problem

$$\chi_{ind(res R)}(g) = \begin{cases} \chi_R(g) + \chi_R(a^{-1}ga) & g \in H \\ 0 & g \notin H \end{cases}$$

Now, let,  $\chi = \chi_{ind(res R)} - \chi_R$ . Then, observe that

$$\chi(g) = \begin{cases} \chi_R(a^{-1}ga) & g \in H \\ -\chi_R(g) & g \notin H \end{cases}$$

But since,

$$\chi_R(a^{-1}ga) = \text{trace}(R_{a^{-1}ga}) = \text{trace}(R_g) = \chi_R(g)$$

We conclude that,  $\chi = \chi_{R'} = \chi_{ind(res R)} - \chi_R$ . This proves that  $ind(res R) = R \oplus R'$ . This completes (b).



Now, we prove (c). Assume  $R$  is irreducible. First we claim that  $R'$  is also irreducible. Indeed,

$$\begin{aligned}\langle \chi_{R'}, \chi_{R'} \rangle &= \frac{1}{|G|} \sum_{g \in G} \chi_{R'}(g) \overline{\chi_{R'}(g)} = \frac{1}{|G|} \left( \sum_{g \in H} \chi_R(g) \overline{\chi_R(g)} + \sum_{g \in aH} (-\chi_R(g)) \overline{(-\chi_R(g))} \right) \\ &= \frac{1}{|G|} \sum_{g \in G} \chi_R(g) \overline{\chi_R(g)} = \langle \chi_R, \chi_R \rangle = 1\end{aligned}$$

. Therefore, our claim is true.

Assume now, that  $R$  and  $R'$  are non-isomorphic. So,  $\langle \chi_R, \chi_{R'} \rangle = 0$  Now,

$$\langle \chi_{resR}, \chi_{resR} \rangle = \langle \chi_{ind(resR)}, \chi_R \rangle = \langle \chi_R + \chi_{R'}, \chi_R \rangle$$

=

$$\langle \chi_R, \chi_R \rangle + \langle \chi_{R'}, \chi_R \rangle = 1$$

. Therefore, we have  $\chi_{resR}$  is irreducible. Observe that we have used Frobenius reciprocity in the first step of the previous equation. We have also used the fact that  $ind(resR) = R \oplus R'$ , to conclude that  $\chi_{ind(resR)} = \chi_R + \chi_{R'}$ .

Next, suppose  $R, R'$  are isomorphic. Then, we have  $\langle \chi_R, \chi_{R'} \rangle = 1$ . From the previous paragraph,

$$\langle \chi_{resR}, \chi_{resR} \rangle = \langle \chi_R, \chi_R \rangle + \langle \chi_{R'}, \chi_R \rangle = 2$$

. Again by the argument in the end of the previous problem, that is, miscellaneous problem number 9, we conclude that  $\chi_{ind(resR)}$  is a sum of two non-isomorphic irreducible representations of  $G$ .

## Result

4 of 4

We have used the properties of induced representation on a group  $G$  from the previous problem, to solve this problem. In this the property called Frobenius reciprocity is the key. Click to see the solution.

## 11. a

To derive the character table of  $S_n$  using induced representations from  $A_n$ , when

[Comment](#)

Step 2 of 18 ^

(a)

Consider the following number:

$$n = 2$$

Order of  $S_3$  is 6

It is known that  $S_3$  has two one dimensional representations, identify this by characters  $\chi_1$  and  $\chi_2$ , namely the one-dimensional representation and the homomorphism of  $S_3$  onto the group  $\{-1, 1\}$  whose kernel is the alternating group  $A_3$ .

There are three classes in  $S_3$  as shown below:

$$C_1 = \{(1)\}, C_2 = \{(12), (23), (13)\}, C_3 = \{(123), (132)\}$$

Therefore,  $S_3$  has three irreducible character and the sum of squares of their degrees must be  $[S_3 : 1] = 6$ .

The third character  $\chi_3$  has degree 2.

[Comment](#)

Step 4 of 18 ^

The character table as shown below:

	$C_1$	$C_2$	$C_3$
$\chi_1$	1	1	1
$\chi_2$	1	-1	1
$\chi_3$	2	$\alpha$	$\beta$

Now determine  $\alpha$  and  $\beta$ , using the orthogonally relations for columns, to have

$$1 \cdot 1 + 1 \cdot 1 + 2 \cdot \beta = 0$$

$$1 \cdot 1 + 1 \cdot (-1) + 2 \cdot \alpha = 0$$

Simplify the above equation, to get  $\alpha$  and  $\beta$ :

$$\beta = -1$$

$$\alpha = 0$$

Therefore, the complete table for  $S_3$  is:

	$C_1$	$C_2$	$C_3$
$\chi_1$	1	1	1
$\chi_2$	1	-1	1
$\chi_3$	2	0	-1

(b)

Consider the following number:

$$n = 4$$

Order of  $S_4$  is 24

The conjugate classes of  $S_4$  are:

$$C_1 = \{(1)\},$$

$$C_2 = \{2\text{-cycles}\},$$

$$C_3 = \{3\text{-cycles}\},$$

$$C_4 = \{4\text{-cycles}\},$$

$$C_5 = \{(12)(34), (13)(24), (14)(23)\}$$

Since,  $[S_4, S_4] = A_4$

There are two one dimensional representation  $\chi_1$  and  $\chi_2$ , where  $\chi_2$  maps the elements of  $A_4$  onto  $+1$  and the elements of  $(12)A_4$  onto  $-1$ .

Then

$$H = \{(1), (12)(34), (14)(23), (13)(24)\}$$

This is a normal subgroup of  $S_4$ , and since  $S_4/H$  is non abelian, so  $S_4/H \cong S_3$  is the only nonabelian group of order 6.

The character  $\chi_3$ , upon composition with the natural homomorphism of  $S_4 \rightarrow S_4/H$  and Third character of degree 2, by evaluate the sum of squares of the degrees to 24, the remaining characters  $\chi_4$  and  $\chi_5$  both have degree 3.

Then,

$$\chi_1^i \overline{\chi_1^i} + \chi_2^i \overline{\chi_2^i} + \chi_3^i \overline{\chi_3^i} + \chi_4^i \overline{\chi_4^i} + \chi_5^i \overline{\chi_5^i} = 0$$

Where,  $i = 2, 3, 4, 5$

This implies that,

$$\chi_4^i + \chi_5^i = 0$$

Where,  $i = 2, 3, 4$

And,

$$\chi_4^5 + \chi_5^5 = -2$$

Therefore, character table is:

	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$\chi_1$	1	1	1	1	1
$\chi_2$	1	-1	1	-1	1
$\chi_3$	2	0	-1	0	2
$\chi_4$	3	$\alpha$	$\beta$	$\gamma$	$\delta$
$\chi_5$	3	$-\alpha$	$-\beta$	$-\gamma$	$-2-\delta$

Now count the number of elements in the different conjugate classes, then

$$h_1 = 1, h_2 = 6, h_3 = 8, h_4 = 6, h_5 = 3$$

And,

$$1 \cdot 1 \cdot 3 + 6 \cdot 1 \cdot \alpha + 8 \cdot 1 \cdot \beta + 6 \cdot 1 \cdot \gamma + 3 \cdot 1 \cdot \delta = 0$$

$$1 \cdot 1 \cdot 3 + 6 \cdot (-1) \cdot \alpha + 8 \cdot 1 \cdot \beta + 6 \cdot (-1) \cdot \gamma + 3 \cdot 1 \cdot \delta = 0$$

$$1 \cdot 2 \cdot 3 + 6 \cdot (0) \cdot \alpha + 8 \cdot (-1) \cdot \beta + 6 \cdot 0 \cdot \gamma + 3 \cdot 2 \cdot \delta = 0$$

Adding the first two equations,

$$6 + 16\beta + 6\delta = 0$$

The last equation written as,

$$6 + -8\beta + 6\delta = 0$$

Simplify the above two equation, to get

$$\beta = 0, \delta = -1$$

Now using the orthogonality relations for column 2 and 4,

$$\alpha\gamma = -1$$

The equation is:

$$\alpha + \gamma = 0$$

Then,  $\alpha = 1, \gamma = -1$

Therefore, the complete table for  $S_4$  as shown below:

	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$\chi_1$	1	1	1	1	1
$\chi_2$	1	-1	1	-1	1
$\chi_3$	2	0	-1	0	2
$\chi_4$	3	1	0	-1	-1
$\chi_5$	3	-1	0	1	-1

(c)

Consider the following number:

$$n = 5$$

Order of  $S_5$  is 120

Find the conjugacy classes of  $S_5$ .

$$C_1 = \{(1)\},$$

$$C_2 = \{(12)\},$$

$$C_3 = \{(1234)\},$$

$$C_4 = \{(12345)\},$$

$$C_5 = \{(12)(34)\}$$

$$C_6 = \{(12)(34)\}$$

The table as shown below:

	1	(12)	(123)	(1234)	(12345)	(12)(34)	(12)(345)
$\chi_1$	1	1	1	1	1	1	1
$\chi_2$	1	-1	1	-1	1	1	-1
$\chi_3$	4	2	1	0	-1	0	-1
$\chi_4$	4	-2	1	0	-1	0	1

Now, find 3 more representations.

The tensor products of two of the representation are:

$$\chi_4 \otimes \chi_4$$

$$\text{But, } \chi_4 \otimes \chi_4 = \chi_5 \oplus \text{sym}^2 \chi_4$$

So, there are two possibilities.

$$\text{Let's compute the first, using } \vartheta_{\chi_4 \chi_4}(g) = \frac{1}{2} \left( \vartheta_{\chi_4}(g)^2 - \vartheta_{\chi_4}(g^2) \right)$$

Then,

$$\begin{aligned} \vartheta_{\chi_4 \chi_4}(g) &= \frac{1}{2} \left( (4, 2, 1, 0, -1, 4, 1)^2 - (4, 4, 1, 0, -1, 4, 1) \right) \\ &= (6, 0, 0, 0, 1, -2, 0) \end{aligned}$$

To compute  $\vartheta_{\chi_4}(g^2)$ :

$$1^2 = 1$$

$$(12)^2 = 1$$

$$(123)^2 = (132)$$

$$(1234)^2 = (13)(24)$$

$$(12345)^2 = (13524)$$

$$((12)(34))^2 = 1$$

$$((12)(345))^2 = (354)$$

The above representation is irreducible.

$$\begin{aligned} \left( \vartheta_{\chi_4}(g^2), \vartheta_{\chi_4}(g^2) \right) &= \frac{1}{120} (36 + 24 \cdot 1 + 15 \cdot 4) \\ &= 1 \end{aligned}$$

There are only two representation left.

Find the dimensions:

$$120 = 1^2 + 1^2 + 4^2 + 4^2 + 6^2 + a^2 + b^2$$

$$a^2 + b^2 = 50$$

So, either one of them has dimension 1 or they both have dimension 5.

There cannot be any representation with dimension 1, the only abelian quotient of  $S_5$  are:

$$S_5/A_5 = \{\pm 1\} \text{ and } S_5/S_5 = \{1\}$$

So, a representation  $\chi_6$ , with character,

$$\vartheta_{\chi_6} = (5, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$$

Now,  $\chi_6 \otimes \chi_2 = \chi_7$  with character

$$\vartheta_{\chi_7} = (5, -\alpha_1, \alpha_2, -\alpha_3, \alpha_4, \alpha_5, -\alpha_6)$$

Check  $\chi_6$  and  $\chi_7$  are orthogonal.

Since  $\chi_6 \neq \chi_7$

This implies that,

$$\alpha_1 = 0, \alpha_3 = 0, \alpha_6 = 0$$

Now it is known that  $\alpha_1 \cdot \alpha_3 \cdot \alpha_6 \neq 0$  and so  $\chi_6 \neq \chi_7$

Therefore, the complete character table for  $S_5$  is:

	1	(12)	(123)	(1234)	(12345)	(12)(34)	(12)(345)
$\chi_1$	1	1	1	1	1	1	1
$\chi_2$	1	-1	1	-1	1	1	-1
$\chi_3$	4	2	1	0	-1	0	-1
$\chi_4$	4	-2	1	0	-1	0	1
$\chi_5$	6	0	0	0	1	-2	0
$\chi_6$	5	1	-1	-1	0	1	1
$\chi_7$	5	-1	-1	1	0	1	-1

12. a

To drive the character table of the dihedral group  $D_n$ , using induced representations from  $C_n$  ;

[Comment](#)

Step 2 of 5 ^

Dihedral group  $D_n$  defined as,

$$D_n = \langle a, b \mid a^n = 1, b^2 = 1, ba = a^{n-1}b \rangle$$

Degree of rotation of element  $a$  is  $\frac{360}{n}$  and all elements  $b, ab, a^2b, \dots, a^{n-1}b$  are reflections.

There is two different pair of conjugacy class,

$$\{a, a^{n-1}\}, \{a^2, a^{n-2}\}$$

If  $n$  is odd, then the last conjugacy class is  $\{a^{(n-1)/2}, a^{(n+1)/2}\}$  and all the reflections are in the same conjugacy class.

So, number of elements in the leading row as shown below,

$$1 + \frac{n-1}{2} + 1 = \frac{n+3}{2}$$

If  $n$  is even, the pairs of elements  $\{a, a^{n-1}\}, \{a^2, a^{n-2}\}, \dots, \{a^{(n-1)/2}, a^{(n+1)/2}\}, \{a^{n/2}\}$  are in different conjugacy classes. The reflections are conjugate either with  $b$  or with  $ab$ .

So, number of elements in the leading row as shown below,

$$1 + \frac{n}{2} + 2 = \frac{n+6}{2}$$



**Consider the case 1:** when  $n$  is odd.

Dihedral group  $D_n$  has two one dimensional representations that map  $a$  to all 1 for all  $i = 0, 1, \dots, n-1$  and  $b$  either to 1 or to  $-1$ .

Then, irreducible representations that map  $a$  to the matrix  $\begin{bmatrix} e^{2k\pi i/n} & 0 \\ 0 & e^{-2k\pi i/n} \end{bmatrix}$  for  $k = 1, \dots, \frac{n-1}{2}$  has  $\frac{n-1}{2}$  two dimensional and  $b$  to  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ .

Thus the traces of these matrices are  $e^{2k\pi i/n} + e^{-2k\pi i/n} = 2 \cos \frac{2k\pi}{n}$  and 0 respectively.

The images of other elements are determined by the image of the generator.

The character tables for  $D_3$  are shown below,

$D_3$	1	a	b
1	1	1	1
$\chi_1$	1	1	-1
$\chi_2$	1	-1	0

**Consider the case2:** when  $n$  is even.

Dihedral group  $D_n$  has four one dimensional representations that map  $a$  to  $\pm 1$  and  $b$  to  $\pm 1$  in all four possible combinations.

Then, representations that map  $a$  to  $\begin{bmatrix} e^{2k\pi i/n} & 0 \\ 0 & e^{-2k\pi i/n} \end{bmatrix}$  for  $k = 1, \dots, \frac{n-2}{2}$  has two  $\frac{n}{2}-1$  dimensional and  $b$  to  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ .

The tables for  $D_2$  and  $D_4$  are shown below,

$D_2$	1	a	b	ab
1	1	1	1	1
$\chi_1$	1	-1	1	-1
$\chi_2$	1	1	-1	-1
$\chi_3$	1	-1	-1	1

And,

$D_4$	1	a	a <sup>2</sup>	b	ab
1	1	1	1	1	1
$\chi_1$	1	-1	1	1	-1
$\chi_2$	1	-1	1	-1	1
$\chi_3$	1	1	1	-1	-1
$\chi_4$	2	0	-2	0	0

13. a

**Given:**  $G$  is a finite subgroup of  $GL_n(\mathbb{C})$  such that

$$\sum_{g \in G} tr(g) = 0.$$

**To Prove:**  $\sum_{g \in G} g = 0$ .

**Proof:** Let us consider

$$X = \sum_{g \in G} g.$$

We will propose to show that  $X$  is a zero matrix. Let us consider  $G$  has  $n$  elements and assume that

$$G = \{g_1, g_2, \dots, g_n\}.$$

Now consider  $n$  bijections  $\rho_i$ , where  $1 \leq i \leq n$  from  $G$  to itself by the assignment

$$\rho_i(X) = g_i X.$$

Now notice that

$$\begin{aligned} X^2 &= \sum_{i=1}^n \left( \sum_{j=1}^n g_j \right) g_i \\ &= \sum_{i=1}^n \sum_{j=1}^n g_j \\ &= \sum_{i=1}^n X \\ &= nX. \end{aligned}$$

Now consider the matrix

$$M = \frac{1}{n}X.$$

Then note that

$$M^2 = \frac{1}{n^2}X^2 = \frac{1}{n^2}(nX) = \frac{1}{n}X = M.$$

Therefore  $M$  is an idempotent matrix.

Now notice that  $\text{GL}_n(\mathbb{C})$  is a group of characteristic zero and we know that in a characteristic zero matrix the rank of an idempotent equals its trace.

Therefore we have

$$\begin{aligned} \text{rank}(M) &= \text{trace}\left(\frac{1}{n}X\right) \\ &= \frac{1}{n}\text{trace}(X) \\ &= \frac{1}{n} \sum_{i=1}^n \text{trace}(g_i) \\ &= 0. \end{aligned}$$

Therefore  $M = 0$ .

Since

$$M = \frac{1}{n}X$$

it follows that

$$X = 0, \text{ i.e. } \sum_{g \in G} g = 0.$$

This completes the proof.

## Result

3 of 3

Considering  $X = \sum_{g \in G} g$  and using the given criterion  $\sum_{g \in G} \text{trace}(g) = 0$  we have proved that  $X$  is a zero matrix.

14. a

Let  $\rho: G \rightarrow GL(V)$  be a representation of  $G$ . Assume that  $\chi$  is its character. Let  $1_G$  denote the trivial representation. It is given that  $\rho(g)$  has one eigen-value 1, for every  $g \in G$ . Suppose,  $|G| = n$ . Now, Let  $\lambda_i$  be the other eigen-value of  $\rho(g_i)$  for every  $1 \leq i \leq n$ . Then, we have,  $\chi(g_i) = 1 + \lambda_i$ .

Now, suppose that the trivial representation occurs in  $\rho$ . Then, it is clear that  $\rho$  is not irreducible, and hence it is the sum of two irreducible representation. We are done in this case.

Now, assume that the trivial representation doesn't occur in  $\rho$ . Then, we have,

$$\langle \chi, 1_G \rangle = \frac{1}{|G|} \sum_{i=1}^n \chi(g_i) = 0$$

Now, as  $\chi(g_i) = 1 + \lambda_i$ , we conclude that

$\sum_{i=1}^n \lambda_i = -n$ , which means  $|\sum_{i=1}^n \lambda_i| = n$ . Now, see that  $\lambda_i$  are roots of unity, and therefore, from that above equation, we conclude that  $\lambda_1 = \lambda_2 = \dots = \lambda_n = \zeta$ . So, now we have that  $\chi(g_i) = 1 + \zeta$ , for each  $1 \leq i \leq n$ . But, then  $\chi(1) = 2$ , and therefore, we conclude that  $\zeta = 1$ . So  $\chi(g) = 2$ , for every  $g \in G$ . Then  $\langle \chi, \chi \rangle = 4$ , and therefore, we see that  $\chi$  is reducible, and hence the sum of two irreducible representation. This completes the proof.

## Result

2 of 2

We break the problem in two cases, and then solve it in each case. We show that if  $\chi$  is the character of  $\rho$ , then  $\langle \chi, \chi \rangle \neq 1$ , and hence is reducible.

15. a

Let  $\rho: G \rightarrow GL_n(\mathbb{C})$  be an irreducible representation of a finite group  $G$ .

Consider the representation  $\rho: GL_n \rightarrow GL_n(V)$  of  $GL_n$  and also consider the representation of  $G$  is composition  $\sigma \circ \rho$ .

[Comment](#)

Step 2 of 4 ^

(a)

To determine the character of the representation of  $\sigma \circ \rho$  when  $\sigma$  is left multiplication of  $GL_n$  on the space  $V$  of  $n \times n$  matrices;

Suppose that  $\sigma \circ \rho$  is representation of  $G$  then define  $\chi_{\sigma \circ \rho}(g)$  is trace of  $g$  on  $\sigma \circ \rho$  where  $g \in GL_n(\mathbb{C})$ .

It is known that a matrix in Jordan canonical form is conjugate to every matrix in  $GL_n(\mathbb{C})$ .

So, character  $\chi_{\sigma \circ \rho}$  defined in Jordan canonical form.

Thus,  $\chi_{\sigma \circ \rho}$  has diagonal matrix as shown below,

$$\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$$

This is denoted by  $\chi_{\sigma \circ \rho}(x_1, \dots, x_n)$ , so this is called the character of  $\sigma \circ \rho$ .

To decompose  $\sigma \circ \rho$  into irreducible representations;

Suppose that  $V = \mathbb{C}^n$ , the standard representation of  $GL_n(\mathbb{C})$  and this acts on direct product of  $V(\otimes e)$  defined by,

$$g(v_1 \otimes \cdots \otimes v_n) = g(v_1) \otimes \cdots \otimes g(v_n)$$

And from the right,  $\sigma \circ \rho(v_1 \otimes \cdots \otimes v_n) = v_{\sigma \circ \rho(i)} \otimes \cdots \otimes v_{\sigma \circ \rho(e)}$

Thus, the above two actions are commute, so  $V(\otimes e)$  is a representation of  $GL_n(\mathbb{C}) \times S_e$

**Therefore,**  $\sigma \circ \rho(v_1 \otimes \cdots \otimes v_n) = v_{\sigma \circ \rho(i)} \otimes \cdots \otimes v_{\sigma \circ \rho(e)}$  is irreducible representation of  $GL_n(\mathbb{C})$ .

(b)

To determine the character of  $\sigma \circ \rho$  when  $\sigma$  is the operation of conjugation on  $\mathbb{C}^{n \times n}$ ;

Suppose that  $\chi$  be a character of  $\sigma \circ \rho$ .

Define  $\chi \sigma \circ \rho(g)$  is trace of  $g$  on  $\sigma \circ \rho$  where  $g \in GL_n(\mathbb{C})$  and  $\sigma$  is the operation of conjugation on  $\mathbb{C}^{n \times n}$ . So

$$\begin{aligned} \chi \sigma \circ \rho(g) &= \text{tr}(\sigma \circ \rho(g)) \\ &= \text{tr}(\sigma(x_1, \dots, x_n)) \\ &= \text{tr}(\sigma(x_1), \dots, \sigma(x_n)) \end{aligned}$$

Where,  $\sigma(x_1), \dots, \sigma(x_n)$  is the operation of conjugation on  $\mathbb{C}^{n \times n}$ .

This can be represented as,

$$\chi \sigma \circ \rho(g) = \begin{pmatrix} \sigma(x_1) & & \\ & \ddots & \\ & & \sigma(x_n) \end{pmatrix}$$

**Thus,** representation  $\chi \sigma \circ \rho(g)$  is the character of  $\sigma \circ \rho$ .

# Chapter 11

## Section 1

1. a

$$\underline{7 + \sqrt[3]{2}}$$

Let

$$y = 7 + \sqrt[3]{2}$$

Then

$$y - 7 = \sqrt[3]{2}$$

This is equivalent to

$$(y - 7)^3 = 2$$

Expanding,

$$y^3 - 21y^2 + 147y - 343 = 2$$

This is equivalent to

$$y^3 - 21y^2 + 147y - 345 = 0$$

Thus, if we define

$$f(x) = x^3 - 21x^2 + 147x - 345$$

we conclude that  $f(7 + \sqrt[3]{2}) = 0$ , so  $7 + \sqrt[3]{2}$  is algebraic.

$$\underline{\sqrt{3} + \sqrt{-5}}$$

$$\text{Let } y = \sqrt{3} + \sqrt{-5}$$

Then

$$y^2 = 3 - 5 + 2\sqrt{-15}$$

Thus,

$$y^2 + 2 = 2\sqrt{-15}$$

Square it:

$$y^4 + 4y^2 + 4 = -60$$

Thus,

$$y^4 + 4y^2 + 64 = 0$$

Therefore, if we define

$$f(x) = x^4 + 4x^2 + 64$$

we conclude that  $f(\sqrt{3} + \sqrt{-5}) = 0$ , so  $\sqrt{3} + \sqrt{-5}$  is algebraic.

## Result

$$(a) f(x) = x^3 - 21x^2 + 147x - 345$$

$$(b) f(x) = x^4 + 4x^2 + 64$$

2. a

By De Moivre's formula,

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx$$

With  $x = \frac{2\pi}{n}$ , we get

$$(\cos(2\pi/n) + i \sin(2\pi/n))^n = 1$$

Using the Binomial Theorem,

$$(\cos(2\pi/n) + i \sin(2\pi/n))^n = \sum_{k=0}^n \binom{n}{k} \cos(2\pi/n)^{n-k} (i)^k \sin(2\pi/n)^k \quad (1)$$

Now recall the following: for  $l \in \mathbb{Z}$ ,

$$i^{4l+0} = 1$$

$$i^{4l+1} = i$$

$$i^{4l+2} = -1$$

$$i^{4l+3} = -i$$

Thus, we separate the sum in (1) into a sum which has no  $i$  and a sum which has  $i$ :

$$\begin{aligned} & \sum_{k=0}^n \binom{n}{k} \cos(2\pi/n)^{n-k} (i)^k \sin(2\pi/n)^k \\ &= \sum_{\substack{0 \leq k \leq n \\ k \text{ even}}} a_k \binom{n}{k} \cos(2\pi/n)^{n-k} \sin(2\pi/n)^k + \sum_{\substack{0 \leq k \leq n \\ k \text{ odd}}} i b_k \binom{n}{k} \cos(2\pi/n)^{n-k} \sin(2\pi/n)^k \end{aligned}$$

Here  $a_k = 1$  if  $k = 4l$  and  $a_k = -1$  if  $k = 4l + 2$ , for some integer  $l$ . Also,  $b_k = 1$  if  $k = 4l + 1$ , and  $b_k = -1$  if  $k = 4l + 3$ .

So, we have that

$$\sum_{\substack{0 \leq k \leq n \\ k \text{ even}}} a_k \binom{n}{k} \cos(2\pi/n)^{n-k} \sin(2\pi/n)^k + i \sum_{\substack{0 \leq k \leq n \\ k \text{ odd}}} b_k \binom{n}{k} \cos(2\pi/n)^{n-k} \sin(2\pi/n)^k = 1 + i \cdot 0$$

Therefore, we have that

$$\sum_{\substack{0 \leq k \leq n \\ k \text{ even}}} a_k \binom{n}{k} \cos(2\pi/n)^{n-k} \sin(2\pi/n)^k = 1 \quad (2)$$



Now we can use the fact that  $k$  is even in the above sum to conclude that

$$\sin(2\pi/n)^k = (\sin(2\pi/n)^2)^{k/2} = (1 - \cos(2\pi/n)^2)^{k/2}$$

Thus, plugging this into (2), and moving 1 to the left-hand side,

$$\sum_{\substack{0 \leq k \leq n \\ k \text{ even}}} a_k \binom{n}{k} \cos(2\pi/n)^{n-k} (1 - \cos(2\pi/n)^2)^{k/2} - 1 = 0$$

If we set  $x = \cos(2\pi/n)$ , then

$$\sum_{\substack{0 \leq k \leq n \\ k \text{ even}}} a_k \binom{n}{k} x^{n-k} (1 - x^2)^{k/2} - 1 = 0$$

Finally, now it is clear that

$$f(x) = \sum_{\substack{0 \leq k \leq n \\ k \text{ even}}} a_k \binom{n}{k} x^{n-k} (1 - x^2)^{k/2} - 1$$

is a polynomial with integer coefficients (recall that Binomial Coefficients are integers!), and  $f(\cos(2\pi/n)) = 0$ .

## Result

First of all

$$(\cos(2\pi/n) + i \sin(2\pi/n))^n = 1$$

(Why?)

Now use the Binomial Theorem.

## 3. a

First notice that

$$\mathbb{Q} \subseteq \mathbb{Q}[\sqrt{2}, \sqrt{3}]$$

and, since  $\sqrt{2}$  and  $\sqrt{3}$  are in  $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$  which is closed under addition,

$$\sqrt{2} + \sqrt{3} \in \mathbb{Q}[\sqrt{2}, \sqrt{3}]$$

So,  $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$  is a subring of  $\mathbb{C}$  which contains  $\mathbb{Q}$  and  $\sqrt{2} + \sqrt{3}$ . Since  $\mathbb{Q}[\sqrt{2} + \sqrt{3}]$  is the smallest such subring, we must have that

$$\mathbb{Q}[\sqrt{2} + \sqrt{3}] \subseteq \mathbb{Q}[\sqrt{2}, \sqrt{3}] \quad (1)$$

On the other hand, notice that

$$(\sqrt{2} + \sqrt{3})^3 = 2\sqrt{2} + 3\sqrt{3} + 6\sqrt{3} + 9\sqrt{2} = 11\sqrt{2} + 9\sqrt{3}$$

Therefore,

$$\sqrt{2} = \frac{1}{2}((\sqrt{2} + \sqrt{3})^3 - 9(\sqrt{2} + \sqrt{3}))$$

Since  $\mathbb{Q}[\sqrt{2} + \sqrt{3}]$  is closed under addition and multiplication, and  $\frac{1}{2}, (\sqrt{2} + \sqrt{3}) \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$ , we get that  $\sqrt{2} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$ .

Similarly,

$$\sqrt{3} = -\frac{1}{2}((\sqrt{2} + \sqrt{3})^3 - 11(\sqrt{2} + \sqrt{3}))$$

and  $\sqrt{3} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$ .

Thus,  $\mathbb{Q}[\sqrt{2} + \sqrt{3}]$  is a subring of  $\mathbb{C}$  which contains  $\mathbb{Q}$ ,  $\sqrt{2}$ , and  $\sqrt{3}$ . Since  $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$  is the smallest such subring, we conclude that

$$\mathbb{Q}[\sqrt{2}, \sqrt{3}] \subseteq \mathbb{Q}[\sqrt{2} + \sqrt{3}] \quad (2)$$

From (1) and (2) we now conclude that

$$\boxed{\mathbb{Q}[\sqrt{2}, \sqrt{3}] = \mathbb{Q}[\sqrt{2} + \sqrt{3}]}$$

First of all,

$$\mathbb{Z}[\gamma] = \{a_n \gamma^n + \dots + a_1 \gamma + a_0 \mid n \in \mathbb{N} \cup \{0\}, a_i \in \mathbb{Z}\}$$

by discussion on page 323 of the book. Furthermore,

$$\begin{aligned} \gamma &= \sqrt{2} + \sqrt{3} \\ \gamma^2 &= (\sqrt{2} + \sqrt{3})^2 = 5 + 2\sqrt{6} \\ \gamma^3 &= (\sqrt{2} + \sqrt{3})^3 = 11\sqrt{2} + 9\sqrt{3} \\ \gamma^4 &= 49 + 20\sqrt{6} \end{aligned}$$

Now we notice that  $\gamma^4 = 10\gamma^2 - 1$ , so

$$\gamma^4 \in \{a_3 \gamma^3 + a_2 \gamma^2 + a_1 \gamma + a_0 \mid a_i \in \mathbb{Z}\}$$

Similarly,  $\gamma^5 = \gamma \cdot \gamma^4 = 10\gamma^3 - \gamma \in \{a_3 \gamma^3 + a_2 \gamma^2 + a_1 \gamma + a_0 \mid a_i \in \mathbb{Z}\}$ , and by induction we conclude that

$$\gamma^n \in \{a_3 \gamma^3 + a_2 \gamma^2 + a_1 \gamma + a_0 \mid a_i \in \mathbb{Z}\}$$

for all  $n \geq 4$ . Thus,

$$\mathbb{Z}[\gamma] = \{a_3 \gamma^3 + a_2 \gamma^2 + a_1 \gamma + a_0 \mid a_i \in \mathbb{Z}\}$$

Now we will first prove that the set  $(1, \sqrt{2}, \sqrt{3}, \sqrt{6})$  is independent over  $\mathbb{Q}$ . Take any linear relation

$$a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} = 0$$

with  $a, b, c, d \in \mathbb{Q}$ . We must prove that  $a = b = c = d = 0$ . Suppose that it is not true; that is, that at least one of them is nonzero. First of all, we can without loss of generality assume that  $a, b, c, d$  are relatively prime integers.

Write the linear relation as

$$a + d\sqrt{6} = -b\sqrt{2} - c\sqrt{3}$$

Square both sides:

$$a^2 + 6d^2 + 2ad\sqrt{6} = 2b^2 + 3c^2 + 2bc\sqrt{6}$$

Thus,

$$(2ad - 2bc)\sqrt{6} = 2b^2 + 3c^2 - a^2 - 6d^2$$

Suppose that  $2ad - 2bc = 0$ , then

$$\sqrt{6} = \frac{2b^2 + 3c^2 - a^2 - 6d^2}{2ad - 2bc}$$

Since  $a, b, c, d$  are rational, we conclude that  $\sqrt{6} \in \mathbb{Q}$ , which is a contradiction since it is not rational. Thus,

$$2ad - 2bc = 0 \implies ad = bc$$

With that we also obtain that

$$2b^2 + 3c^2 - a^2 - 6d^2 = 0$$

If  $b = 0$ , then  $ad = 0$ , which means that either  $a = 0$  or  $d = 0$ . Suppose that  $d = 0$ . Then

$$3c^2 - a^2 = 0 \implies a^2 = 3c^2$$

This means that 3 divides  $a^2$ , so it also divides  $a$  since 3 is prime. Thus,  $a = 3k$ , for some integer  $k$ , and

$$9k^2 = 3c^2 \implies c^2 = 3k^2$$

This also means that 3 divides  $c^2$  and 3 divides  $c$ . Since 3 trivially divides 0, it also divides  $b$  and  $d$ . Thus, 3 is a common divisor of  $a, b, c, d$  which is impossible since we assumed that they are relatively prime.

Therefore,  $d \neq 0$ . If we suppose that  $a = 0$ , we get  $3c^2 - 6d^2 = 0$ , or  $c^2 = 2d^2$ . Now we show that 2 is a common divisor of  $a, b, c, d$  which is impossible.

Therefore,  $b \neq 0$ . Similar conclusion follows if we assume  $c = 0$ . Therefore,  $c \neq 0$ .

So,  $ad \neq 0$ , so  $d \neq 0$ . Now  $ad = bc$  can be written as

$$\frac{a}{c} = \frac{b}{d} = t$$

Now observe the equation

$$2b^2 + 3c^2 - a^2 - 6d^2 = 0 \implies a^2 + 6d^2 = 2b^2 + 3c^2$$

Since  $b = dt$  and  $a = ct$ , this equation becomes

$$c^2t^2 + 6d^2 = 2d^2t^2 + 3c^2$$

This can be simplified:

$$t^2(c^2 - 2d^2) = 3(c^2 - 2d^2),$$

or

$$(t^2 - 3)(c^2 - 2d^2) = 0$$

So,  $t^2 - 3 = 0$  or  $c^2 - 2d^2 = 0$ . However,

$$t^2 - 3 = 0 \implies t = \pm\sqrt{3},$$

which is impossible since  $t \in \mathbb{Q}$ , and  $\pm\sqrt{3} \notin \mathbb{Q}$ . Thus, we would have

$$c^2 - 2d^2 = 0 \implies \sqrt{2} = \frac{|c|}{|d|}$$

This is also impossible, since  $\frac{|c|}{|d|} \in \mathbb{Q}$  and  $\sqrt{2} \notin \mathbb{Q}$ .

Finally, we obtained a contradiction, which means that we must have that

$$a = b = c = d = 0$$

Now to complete our exercise. Suppose that

$$\sqrt{2} \in \mathbb{Z}[\alpha]$$

Then there exist some integers  $a, b, c, d$  such that

$$\sqrt{2} = a + b(\sqrt{2} + \sqrt{3}) + c(5 + 2\sqrt{6}) + d(11\sqrt{2} + 9\sqrt{3})$$

That is,

$$a + (b + 11d - 1)\sqrt{2} + (b + 9d)\sqrt{3} + 2c\sqrt{6} = 0$$

Since  $(1, \sqrt{2}, \sqrt{3}, \sqrt{6})$  is independent over  $\mathbb{Q}$ , the above equation is equivalent to the system

$$\begin{aligned} a &= 0 \\ b + 11d &= 1 \\ b + 9d &= 0 \\ 2c &= 0 \end{aligned}$$

Now subtract the third equation from the second:

$$2d = 1$$

However, this is impossible, since we would have  $d = \frac{1}{2} \notin \mathbb{Z}$ . Thus,

$$\sqrt{2} \notin \mathbb{Z}[\gamma],$$

and

$$\boxed{\mathbb{Z}[\sqrt{2}, \sqrt{3}] \neq \mathbb{Z}[\sqrt{2} + \sqrt{6}]}$$

## Result

6

The first equality holds, the second does not.

Hint for the first equality: show that  $\gamma \in \mathbb{Q}[\alpha, \beta]$ , and  $\alpha, \beta \in \mathbb{Q}[\gamma]$ . Why does the equality hold from this?

Hint for the second part: First show that  $\mathbb{Z}[\gamma] = \{a + b\gamma + c\gamma^2 + d\gamma^3 \mid a, b, c, d \in \mathbb{Z}\}$ . Then show that  $(1, \sqrt{2}, \sqrt{3}, \sqrt{6})$  is independent over  $\mathbb{Q}$ . From this, show that  $\sqrt{2} \notin \mathbb{Z}[\gamma]$ .

4. a

First of all, by discussion on the page 323,

$$\mathbb{Z}[\alpha] = \{a_n \alpha^n + \dots + a_1 \alpha + a_0 \mid n \in \mathbb{N} \cup \{0\}, a_i \in \mathbb{Z}\}$$

Furthermore, recall that, for any  $l \in \mathbb{Z}$ ,

$$\begin{aligned} i^{4l} &= 1 \\ i^{4l+1} &= i \\ i^{4l+2} &= -1 \\ i^{4l+3} &= -i \end{aligned}$$

Therefore,

$$\beta \in \mathbb{Z}[\alpha] \iff \beta = a + bi,$$

where  $a, b$  are of the form

$$\frac{c_n}{2^n} + \frac{c_{n-2}}{2^{n-2}} + \dots + c_0$$

or

$$\frac{c_n}{2^n} + \frac{c_{n-2}}{2^{n-2}} + \dots + \frac{c_1}{2},$$

with  $c_i \in \mathbb{Z}$ . Furthermore, by writing  $c_k = 2d_{k-1} + d_k$ , for  $k \geq 1$ , we get that  $a, b$  are of the form

$$\frac{d_n}{2^n} + \frac{d_{n-1}}{2^{n-1}} + \dots + \frac{d_1}{2} + d_0$$

This means that

$$a, b \in \mathbb{Z}[1/2]$$

Now to solve the exercise. Let  $z = x + yi \in \mathbb{C}$  and  $\varepsilon > 0$  be arbitrarily taken. Then  $x, y \in \mathbb{R}$ . Since  $\mathbb{Z}[1/2]$  is dense in  $\mathbb{R}$ , there exist  $a, b \in \mathbb{Z}[1/2]$  which are  $\frac{\varepsilon}{\sqrt{2}}$ -close to  $x$  and  $y$ , respectively. This means that  $|a - x| < \frac{\varepsilon}{\sqrt{2}}$  and  $|b - y| < \frac{\varepsilon}{\sqrt{2}}$ . Now, by digression from the start of the proof, we conclude that  $a + bi \in \mathbb{Z}[\alpha]$ .

Furthermore, now we conclude that

$$|z - (a + bi)| = \sqrt{|x - a|^2 + |y - b|^2} < \sqrt{\frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2}} = \varepsilon$$

Since  $\varepsilon$  and  $z$  were taken arbitrarily, we conclude that  $\mathbb{Z}[\alpha]$  is dense in the complex plane.

## Result

Hint: you can show that

$$\mathbb{Z}[\alpha] = \mathbb{Z}[1/2] + i\mathbb{Z}[1/2]$$

That is, that  $\beta \in \mathbb{Z}[\alpha]$  if and only if  $\beta = a + bi$ , where  $a, b \in \mathbb{Z}[1/2]$ .

5. a



Let  $S \subseteq \mathbb{R}$  be a subring of  $\mathbb{R}$ . Then, by definition,  $1 \in S$ . This also means that, for every  $n \in \mathbb{N}$ ,

$$n = \underbrace{1 + 1 + \dots + 1}_{n \text{ times}} \in S,$$

since  $S$  is closed under addition.

Furthermore,  $S$  is closed under subtraction, so

$$0 = 1 - 1 \in S$$

and

$$-1 = 0 - 1 \in S$$

Thus,

$$-n = \underbrace{(-1) + (-1) + \dots + (-1)}_{n \text{ times}} \in S$$

Therefore,

$$\mathbb{Z} \subseteq S,$$

and  $\mathbb{Z}$  is clearly discrete subset of  $\mathbb{R}$ .

Now suppose that some noninteger  $\alpha \in S$ . We can write

$$\alpha = m + \beta,$$

where  $m \in \mathbb{Z}$  and  $0 < \beta < 1$ . Since  $S$  is closed under multiplication,

$$\beta^n = \underbrace{\beta \cdot \beta \cdots \beta}_{n \text{ times}} \in S,$$

for every  $n \in \mathbb{N}$ .

Since  $|\beta| < 1$ , we conclude that

$$\lim_{n \rightarrow \infty} \beta^n = 0$$

Therefore, for every  $\varepsilon > 0$ , there exists some  $s \in S$  such that  $|s| < \varepsilon$ , so  $S$  is not a discrete subset of  $\mathbb{R}$ .

Thus, there cannot be any noninteger element in  $S$ .

## Result

Only  $\mathbb{Z}$ .

6. a

(a)

We will prove that this subset is closed under addition, subtraction, and multiplication, and that it contains 1.

Closed under addition?

Let  $a/b, c/d \in S$ . Then  $b$  and  $d$  are not divisible by 3. By definition,

$$a/b + c/d = \frac{a \cdot \frac{d}{\text{lcm}(b,d)} + c \cdot \frac{b}{\text{lcm}(b,d)}}{\text{lcm}(b,d)},$$

where  $\text{lcm}$  is the *least common multiple*. Suppose that 3 divides  $\text{lcm}(b, d)$ . Since  $bd$  is a common multiple of  $b$  and  $d$ , by definition of  $\text{lcm}$  we have that  $\text{lcm}(b, d)$  divides  $bd$ . Thus, 3 divides  $bd$ . Since 3 is prime, this means that 3 divides either  $b$  or  $d$ . This is a contradiction, since  $b$  and  $d$  are **not** divisible by 3. Thus, 3 does not divide  $\text{lcm}(b, d)$ , so  $a/b + c/d \in S$ .

Closed under subtraction?

Let  $a/b, c/d \in S$ . Since

$$a/b - c/d = a/b + (-c)/d,$$

and clearly  $(-c)/d \in S$ , closure under subtraction follows from the fact that  $S$  is closed under addition.

Closed under multiplication?

Let  $a/b, c/d \in S$ . Then

$$(a/b) \cdot (c/d) = \frac{ac}{bd}$$

Suppose that 3 divides  $bd$ . Then 3 divides either  $b$  or  $d$ . Both is a contradiction since  $a/b, c/d \in S$ . Thus, 3 does not divide  $bd$ , and  $(a/b) \cdot (c/d) \in S$ .

1 ∈ S?

We write

$$1 = 1/1,$$

so  $1 \in S$ , since 3 does not divide 1.

(b)

We will prove that  $S$  is not closed under multiplication. To prove that, we will first prove that  $S$  is linearly independent over  $\mathbb{R}$ .

First of all, since  $\cos(-nt) = \cos nt$  and  $\sin(-nt) = -\sin nt$ ,  $S$  is the set of all linear combinations with integer coefficients of the functions  $\{1, \sin nt, \cos nt \mid n \in \mathbb{N}\}$ .

Take any linear relation

$$c_n v_n + c_{n-1} v_{n-1} + \dots + c_1 v_1 + c_0 v_0 = 0, \quad (1)$$

where  $v_i \in \{1, \sin nt, \cos nt \mid n \in \mathbb{N}\}$ ,  $c_i \in \mathbb{R}$ . Using the trigonometric identity  $\sin x \sin y = \frac{\cos(x-y) - \cos(x+y)}{2}$ , for  $n \neq m$  we have that

$$\begin{aligned} \int_0^{2\pi} \sin nt \sin mt dt &= \frac{1}{2} \int_0^{2\pi} \cos(n-m)t dt + \frac{1}{2} \int_0^{2\pi} \cos(n+m)t dt \\ &= \frac{1}{2(n-m)} \sin(n-m)t \Big|_0^{2\pi} + \frac{1}{2(n+m)} \sin(n+m)t \Big|_0^{2\pi} \\ &= 0 \end{aligned}$$

Similarly,

$$\int_0^{2\pi} \cos nt \cos mtdt = 0$$

$$\int_0^{2\pi} \sin nt \cos mtdt = 0$$

$$\int_0^{2\pi} \cos nt \sin mtdt = 0$$

Also,

$$\int_0^{2\pi} \cos ntdt = \frac{1}{n} \sin nt \Big|_0^{2\pi} = 0$$

and

$$\int_0^{2\pi} \sin ntdt = -\frac{1}{n} \cos nt \Big|_0^{2\pi} = 0$$

However,

$$\int_0^{2\pi} 1dt = 2\pi$$

$$\int_0^{2\pi} (\sin nt)^2 dt = \int_0^{2\pi} \left( \frac{1 - \cos 2nt}{2} \right) dt = \int_0^{2\pi} \frac{dt}{2} - \frac{1}{2} \int_0^{2\pi} \cos 2ntdt = \pi$$

$$\int_0^{2\pi} (\cos nt)^2 dt = \int_0^{2\pi} \left( \frac{1 + \cos 2nt}{2} \right) dt = \int_0^{2\pi} \frac{dt}{2} + \frac{1}{2} \int_0^{2\pi} \cos 2ntdt = \pi$$

Now we return to (I). Fix some  $i \in \{0, 1, \dots, n\}$ . We multiply (I) by  $v_i$  to get

$$c_n v_n v_i + \dots + c_i v_i^2 + \dots + c_0 v_0 v_i = 0$$

Use  $\int_0^{2\pi} dt$ :

$$c_n \int_0^{2\pi} v_n v_i dt + \dots + c_i \int_0^{2\pi} v_i^2 dt + \dots + c_0 \int_0^{2\pi} v_0 v_i dt = 0$$

Now we conclude that  $\int_0^{2\pi} v_j v_i dt = 0$ , for  $j \neq i$ . Thus, the above equality becomes

$$c_i \int_0^{2\pi} v_i^2 dt = 0$$

However,  $\int_0^{2\pi} v_i^2 dt \neq 0$ , so

$$\boxed{c_i = 0}$$

Since  $i \in \{0, 1, \dots, n\}$  was taken arbitrarily,

$$\boxed{c_0 = c_1 = \dots = c_n = 0}$$

Thus, the set  $\{1, \cos nt, \sin nt \mid n \in \mathbb{N}\}$  is linearly independent over  $\mathbb{R}$ .

This gives us a powerful tool. If  $x = c_n v_n + \dots + c_1 v_1$  and  $x = d_n v_n + \dots + d_1 v_1$ , for  $v_i \in \{1, \cos nt, \sin nt \mid n \in \mathbb{N}\}$  and  $c_i, d_i \in \mathbb{R}$ , then

$$(c_n - d_n)v_n + \dots + (c_1 - d_1)v_1 = 0,$$

and  $c_i - d_i = 0$ ; that is,  $c_i = d_i$ , for all  $i = 1, \dots, n$ . This means that every vector from the span of  $\{1, \cos nt, \sin nt \mid n \in \mathbb{N}\}$  can be written as a linear combination of elements of  $\{1, \cos nt, \sin nt \mid n \in \mathbb{N}\}$  in only one way.

Finally, we see that

$$\sin t \cos t = \frac{1}{2} \sin 2t$$

By the digression from before, this is the only way to write  $\sin t \cos t$  as a linear combination of elements of  $\{1, \cos nt, \sin nt \mid n \in \mathbb{N}\}$ . However,  $\frac{1}{2} \notin \mathbb{Z}$ , so  $\sin t \cos t \notin S$ , while  $\sin t$  and  $\cos t$  are from  $S$ ! This means that  $S$  is not closed under multiplication, so it cannot be a subring.

## Result

(a) Yes.

(b) No. Hint:  $\cos t \sin t$  is not in  $S$ .

7. a

(a)

This is a ring. We check all properties from Definition 11.1.3.

+ makes  $R$  into an abelian group.

1. The fact that  $A + B \in R$  is trivial.
2. We need to check that  $+$  is associative. By  $S'$  we denote the set  $U - S$ .

$$\begin{aligned} & (A + B) + C \\ &= ((A + B) \cup C) - ((A + B) \cap C) \\ &= (((A \cup B) - (A \cap B)) \cup C) - (((A \cup B) - (A \cap B)) \cap C) \\ &= (((A \cup B) \cap (A' \cup B')) \cup C) \cap ((A \cup B) \cap (A' \cup B')) \cap C' \\ &= (((A \cap A') \cup (A \cap B') \cup (B \cap A') \cup (B \cap B')) \cup C) \cap ((A' \cap B') \cup (A \cap B) \cup C') \\ &= ((A \cap B') \cup (B \cap A') \cup C) \cap ((A' \cap B') \cup (A \cap B) \cup C') \\ &= (A \cap B \cap C) \cup (A \cap B' \cap C') \cup (A' \cap B \cap C') \cup (A' \cap B' \cap C) \end{aligned}$$

Similarly,

$$A + (B + C) = (A \cap B \cap C) \cup (A \cap B' \cap C') \cup (A' \cap B \cap C') \cup (A' \cap B' \cap C)$$

Thus,

$$(A + B) + C = A + (B + C)$$

3. The empty set  $\emptyset$  will be the additive inverse:

$$A + \emptyset = (A \cup \emptyset) - (A \cap \emptyset) = A - \emptyset = A$$

$$\emptyset + A = (\emptyset \cup A) - (\emptyset \cap A) = A - \emptyset = A$$

4. Let  $A$  be some subset of  $U$ . Then

$$A + A = (A \cup A) - (A \cap A) = A - A = \emptyset = 0$$

Thus, each element has an additive inverse; itself.

5. To prove that  $+$  is commutative, let  $A, B$  be two subsets of  $U$ . Then,

$$A + B = (A \cup B) - (A \cap B) = (B \cup A) - (B \cap A) = B + A$$

since  $\cup$  and  $\cap$  are commutative operations.

$\cdot$  is associative, commutative, and has an identity 1.

1. For  $A, B \in R$ ,  $A \cdot B \in R$  is trivial.

2. Let  $A, B, C \in R$ . Then

$$(A \cdot B) \cdot C = (A \cap B) \cdot C = (A \cap B) \cap C = A \cap (B \cap C) = A \cdot (B \cdot C)$$

since  $\cap$  is an associative operation.

3. Notice that, for every  $A \in R$ ,

$$A \cdot U = A \cap U = A$$

$$U \cdot A = U \cap A = A$$

since  $A \subseteq U$ . Thus,  $U$  is a multiplicative identity.

4. Let  $A, B$  be two subsets of  $U$ . Then

$$A \cdot B = A \cap B = B \cap A = B \cdot A$$

since  $\cap$  is a commutative operation.

Distributive law.

Let  $A, B, C \in R$ . Then

$$\begin{aligned} (A + B) \cdot C &= ((A \cup B) - (A \cap B)) \cap C \\ &= ((A \cup B) \cap (A' \cup B')) \cap C \\ &= ((A \cap B') \cup (A' \cap B)) \cap C \\ &= (A \cap B' \cap C) \cup (A' \cap B \cap C) \end{aligned}$$

On the other hand,

$$\begin{aligned} (A \cdot C) + (B \cdot C) &= (A \cap C) + (B \cap C) \\ &= ((A \cap C) \cup (B \cap C)) - ((A \cap C) \cap (B \cap C)) \\ &= ((A \cup B) \cap (A \cup C) \cap (B \cup C) \cap C) \cap (A \cap B \cap C)' \\ &= ((A \cup B) \cap (A \cup C) \cap (B \cup C) \cap C) \cap (A' \cup B' \cup C') \\ &= (A \cap B' \cap C) \cup (A' \cap B \cap C) \end{aligned}$$

Therefore, the distributive law holds.

(b)

This is not a ring. We still check all properties from Definition 11.1.3.  
+ makes  $R$  into an abelian group.

1. The fact that  $f + g \in R$  is trivial, since the sum of two continuous functions is a continuous function.
2. Let  $f, g, h \in R$ . Then

$$[[f + g] + h](x) = [f + g](x) + h(x) = f(x) + g(x) + h(x)$$

$$[f + [g + h]](x) = f(x) + [g + h](x) = f(x) + g(x) + h(x)$$

Therefore,  $[[f + g] + h](x) = [f + [g + h]](x)$  for every  $x \in \mathbb{R}$ , so  $[f + g] + h = f + [g + h]$ , as required.

3. Let  $n : \mathbb{R} \rightarrow \mathbb{R}$ ,  $n(x) = 0$ . Then

$$[f + n](x) = f(x) + n(x) = f(x)$$

$$[n + f](x) = n(x) + f(x) = f(x)$$

Therefore, for every  $x \in \mathbb{R}$  we have that

$$[f + n](x) = [n + f](x) = f(x),$$

so

$$f + n = n + f = f$$

Therefore,  $n$  is an additive identity.

4. Let  $f \in R$ . Define  $g(x) = -f(x)$ .  $g$  is continuous, and

$$[f + g](x) = [g + f](x) = 0 = n(x)$$

for every  $x \in \mathbb{R}$ , so

$$f + g = g + f = n$$

and  $g = -f$  is the additive inverse of  $f$  in  $R$ .

5. Let  $f, g \in R$ , then

$$[f + g](x) = f(x) + g(x) = g(x) + f(x) = [g + f](x),$$

since the addition of real numbers is commutative. Thus  $f + g = g + f$ .

$\cdot$  is associative, and has an identity 1, but is not commutative.

1. The fact that  $f \cdot g \in R$  is trivial, since the composition of two continuous functions is a continuous function.
2. Let  $f, g, h \in R$ . Then

$$(f \circ g) \circ h = f \circ (g \circ h)$$

by properties of composition. Thus,  $\cdot$  is associative.

3. Let  $id : \mathbb{R} \rightarrow \mathbb{R}$ ,  $id(x) = x$ . Then

$$(f \cdot id)(x) = f(id(x)) = f(x)$$

$$(id \cdot f)(x) = id(f(x)) = f(x)$$

Therefore, for every  $x \in \mathbb{R}$  we have that

$$(f \cdot id)(x) = (id \cdot f)(x) = f(x),$$

so

$$f \cdot id = id \cdot f = f$$

Therefore,  $id$  is a multiplicative identity.



4. This is not a commutative law of composition! For example, let  $f(x) = x^2$ ,  $g(x) = x + 1$ . Then

$$(f \cdot g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2 = x^2 + 2x + 1$$

$$(g \cdot f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

However,

$$x^2 + 2x + 1 \neq x^2 + 1,$$

so  $f \cdot g \neq g \cdot f$ .

#### Distributive law.

This also does not hold in its entirety. The right distributive law holds. Let  $f, g, h \in R$ . Then

$$[(f + g) \cdot h](x) = [f + g](h(x)) = f(h(x)) + g(h(x)) = [f \cdot h](x) + [g \cdot h](x) = [(f \cdot h) + (g \cdot h)](x)$$

Thus,

$$(f + g) \cdot h = (f \cdot h) + (g \cdot h)$$

On the other hand, the left distributive law does not hold. Let  $f(x) = x^2$ ,  $g(x) = h(x) = x$ . Then

$$[f \cdot (g + h)](x) = f([g + h](x)) = f(g(x) + h(x)) = f(x + x) = f(2x) = (2x)^2 = 4x^2$$

On the other hand,

$$[(f \cdot g) + (f \cdot h)](x) = [f \cdot g](x) + [f \cdot h](x) = f(g(x)) + f(h(x)) = f(x) + f(x) = 2x^2$$

Since  $2x^2 \neq 4x^2$ , we conclude that

$$f \cdot (g + h) \neq (f \cdot g) + (f \cdot h)$$

#### **Result**

(a) This is a ring.

(b) This is not a ring. The multiplication is not commutative, and the distributive law does not hold.

8. a

We will first prove that following statement which will help us greatly: if  $x$  is a unit, then the equality  $xy = 0$  implies that  $y = 0$ . This is seen almost immediately; just multiply the equality  $xy = 0$  by  $x^{-1}$  to get  $y = 0$ .

**(a)**

First of all, 0 is clearly not a unit.

$1 \cdot 1 = 1$ , so 1 is a unit.

$2 \cdot 6 = 12 \neq 0$ , and  $6 \neq 0$ , so 2 cannot be unit.

$3 \cdot 4 = 12 \neq 0$ , and  $4 \neq 0$ , so 3 cannot be unit.

$4 \cdot 3 = 12 \neq 0$ , and  $3 \neq 0$ , so 4 cannot be unit.

$5 \cdot 5 = 25 = 1$ , so 5 is a unit.

$6 \cdot 2 = 12 \neq 0$ , and  $2 \neq 0$ , so 6 cannot be a unit.

$7 \cdot 7 = 49 = 1$ , so 7 is a unit.

$8 \cdot 3 = 24 \neq 0$ , and  $3 \neq 0$ , so 8 cannot be a unit.

$9 \cdot 4 = 36 \neq 0$ , and  $4 \neq 0$ , so 9 cannot be a unit.

$10 \cdot 6 = 60 \neq 0$ , and  $6 \neq 0$ , so 10 cannot be a unit.

$11 \cdot 11 = 121 = 1$ , so 11 is a unit.

So, all units are

1, 5, 7, 11

**(b)**

First of all, 0 is clearly not a unit.

$1 \cdot 1 = 1$ , so 1 is a unit.

$2 \cdot 4 = 8 \neq 0$ , and  $4 \neq 0$ , so 2 cannot be unit.

$3 \cdot 3 = 9 = 1$ , so 3 is a unit.

$4 \cdot 2 = 8 \neq 0$ , and  $2 \neq 0$ , so 4 cannot be unit.

$5 \cdot 5 = 25 = 1$ , so 5 is a unit.

$6 \cdot 4 = 24 \neq 0$ , and  $4 \neq 0$ , so 6 cannot be a unit.

$7 \cdot 7 = 49 = 1$ , so 7 is a unit.

So, all units are

1, 3, 5, 7

(c)

We will prove that  $a$  is a unit if and only if  $a$  and  $n$  are relatively prime.

Suppose that  $a$  is a unit. Then there exists some  $b \in \mathbb{Z}/n\mathbb{Z}$  such that

$$ab = 1$$

This means that

$$ab \equiv 1 \text{ modulo } n$$

Thus,

$$ab - 1 = nk,$$

for some integer  $k$ . Furthermore, we now have that

$$ab + n(-k) = 1$$

Since the greatest common divisor is the smallest positive integer which can be written in the above way, and 1 is the smallest positive integer of them all, we conclude that the greatest common divisor of  $a$  and  $n$  is 1; that is,  $a$  and  $n$  are relatively prime.

The converse is analogous; suppose that  $a$  and  $n$  are relatively prime. Then there exist integers  $k, l$  such that

$$ak + nl = 1$$

That is, we have that

$$ak - 1 = nl,$$

so  $ak \equiv 1 \text{ modulo } n$ , or  $ak = 1$  in  $\mathbb{Z}/n\mathbb{Z}$ . So,  $k = a^{-1}$ , and  $a$  is a unit.

---

## Result

(a) 1, 5, 7, 11

(b) 1, 3, 5, 7

(c)  $a$  is a unit if and only if  $a$  and  $n$  are relatively prime.

9. a

We first prove the following statement:

$$-x = (-1)x,$$

for all  $x \in R$ . Truly,

$$(-1)x + x = (-1)x + 1 \cdot x = [(-1) + 1]x = 0x = 0$$

$$x + (-1)x = 1 \cdot x + (-1)x = [1 + (-1)]x = 0x = 0$$

Therefore,

$$(-1)x + x = x + (-1)x = 0$$

Thus,  $(-1)x = -x$ , the additive inverse of  $x$ .

Now let  $a, b \in R$ . Then  $a + b \in R$ , and  $-(b + a) \in R$ . Moreover,

$$-(b + a) = (-1)(b + a) \stackrel{(*)}{=} (-1)b + (-1)a = -b + (-a),$$

where we used the distributive law in  $(*)$ . Now,

$$\begin{aligned}(a + b) + (-(b + a)) &= (a + b) + (-b + (-a)) \\ &\stackrel{(**)}{=} a + (b + (-b)) + (-a) \\ &= a + 0 + (-a) \\ &= a + (-a) \\ &= 0,\end{aligned}$$

where we used the associative property of addition in  $(**)$ . Similarly, excluding the details here,

$$-(b + a) + (a + b) = -b + (-a) + a + b = 0$$

Therefore,

$$(a + b) + (-(b + a)) = -(b + a) + (a + b) = 0,$$

hence

$$-(a + b) = -(b + a),$$

that is,

$$(-1)(a + b) = (-1)(b + a)$$

Multiplying the above equality by  $-1$  yields

$$a + b = b + a,$$

as required.

## Result

Hint: show that  $-(b + a) = (-1)(b + a)$  and that  $-(b + a) = -(a + b)$ .

## Section 2

1. a

By definition, this means that there exists a polynomial  $q(x)$  such that

$$x^4 + 3x^3 + x^2 + 7x + 5 = (x^2 + x + 1)q(x) \quad (1)$$

Notice that the product of  $x^2 + x + 1$  and  $q(x)$  will be of degree  $2 + \deg q(x)$ , where  $\deg q(x)$  is the degree of  $q(x)$ . Since the polynomial on the left-hand side is of degree 4, we conclude that  $q(x)$  is of degree 2. Thus,

$$q(x) = ax^2 + bx + c$$

Plugging this into (1), we obtain the equality

$$x^4 + 3x^3 + x^2 + 7x + 5 = (x^2 + x + 1)(ax^2 + bx + c)$$

Furthermore,

$$(x^2 + x + 1)(ax^2 + bx + c) = ax^4 + (a + b)x^3 + (a + b + c)x^2 + (b + c)x + c$$

So we obtain a polynomial equation

$$x^4 + 3x^3 + x^2 + 7x + 5 = ax^4 + (a + b)x^3 + (a + b + c)x^2 + (b + c)x + c$$

The corresponding coefficients must be equal, so we obtain a system of equations

$$\begin{aligned} a &= 1 \\ a + b &= 3 \\ a + b + c &= 1 \\ b + c &= 7 \\ c &= 5 \end{aligned}$$

Thus, we immediately get  $a = 1$ ,  $c = 5$ . Furthermore, from the second and fourth equation we get that  $b = 2$ . Plugging all this into the third equation:

$$1 + 2 + 5 = 1 \iff 7 = 0$$

This means that in  $\mathbb{Z}/n\mathbb{Z}$ , the equality  $7 = 0$  must hold, which means that  $n = 7$ .

## Result

2 of 2

$$n = 7$$

## 2. a

We define the operations the same way as with polynomials. Let

$$f(t) = a_0 + a_1t + a_2t^2 + \dots$$

$$g(t) = b_0 + b_1t + b_2t^2 + \dots$$

Then

$$f(t) + g(t) = (a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \dots = \sum_{n=0}^{\infty} (a_n + b_n)t^n$$

$$f(t)g(t) = \sum_{n=0}^{\infty} c_nt^n, \quad c_n = \sum_{k=0}^n a_kb_{n-k}$$

+ makes  $F[[t]]$  into an abelian group.

Closure. If  $f(t), g(t) \in F[[t]]$ , then clearly  $f(t) + g(t) \in F[[t]]$ .

Associativity. Let  $f(t), g(t), h(t) \in F[[t]]$ ,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n, \quad g(t) = \sum_{n=0}^{\infty} b_n t^n, \quad h(t) = \sum_{n=0}^{\infty} c_n t^n$$

Then

$$\begin{aligned} (f(t) + g(t)) + h(t) &= \left( \sum_{n=0}^{\infty} (a_n + b_n) t^n \right) + h(t) \\ &= \sum_{n=0}^{\infty} ((a_n + b_n) + c_n) t^n \\ &\stackrel{(*)}{=} \sum_{n=0}^{\infty} (a_n + b_n + c_n) t^n \end{aligned}$$

(In  $(*)$  we used that  $F$  is a field, so the addition of elements of  $F$  is associative.)

Similarly,

$$\begin{aligned} f(t) + (g(t) + h(t)) &= f(t) + \left( \sum_{n=0}^{\infty} (b_n + c_n) t^n \right) \\ &= \sum_{n=0}^{\infty} (a_n + (b_n + c_n)) t^n \\ &\stackrel{(*)}{=} \sum_{n=0}^{\infty} (a_n + b_n + c_n) t^n \end{aligned}$$

(Again, in  $(*)$  we used that  $F$  is a field, so the addition of elements of  $F$  is associative.)

Finally, we conclude that

$$(f(t) + g(t)) + h(t) = f(t) + (g(t) + h(t))$$

Identity. Define

$$n(t) = 0 = \sum_{n=0}^{\infty} 0 \cdot t^n \in F[[t]]$$

( $0 \in F$  since  $F$  is a field).

Then, for every  $f(t) \in F[[t]]$ ,  $f(t) = \sum_{n=0}^{\infty} a_n t^n$ , we have that

$$\begin{aligned} f(t) + n(t) &= \sum_{n=0}^{\infty} (a_n + 0) t^n = \sum_{n=0}^{\infty} a_n t^n = f(t) \\ n(t) + f(t) &= \sum_{n=0}^{\infty} (0 + a_n) t^n = \sum_{n=0}^{\infty} a_n t^n = f(t) \end{aligned}$$

Thus,

$$f(t) + n(t) = n(t) + f(t) = f(t)$$



for every  $f(t) \in F[[t]]$ . Therefore,  $n(t)$  is the additive identity.

Additive inverse. Let  $f(t) = \sum_{n=0}^{\infty} a_n t^n$ . Define

$$g(t) = \sum_{n=0}^{\infty} (-a_n) t^n \in F[[t]]$$

( $-a_n$  exists because  $F$  is a field).

Then

$$\begin{aligned} f(t) + g(t) &= \sum_{n=0}^{\infty} (a_n + (-a_n)) t^n = \sum_{n=0}^{\infty} 0 \cdot t^n = n(t) \\ g(t) + f(t) &= \sum_{n=0}^{\infty} (-a_n + a_n) t^n = \sum_{n=0}^{\infty} 0 \cdot t^n = n(t) \end{aligned}$$

Thus,  $g(t)$  is the additive inverse of  $f(t)$ .

Commutativity. Let  $f(t), g(t) \in F[[t]]$ ,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n, \quad g(t) = \sum_{n=0}^{\infty} b_n t^n$$

Then

$$f(t) + g(t) = \sum_{n=0}^{\infty} (a_n + b_n) t^n \stackrel{(**)}{=} \sum_{n=0}^{\infty} (b_n + a_n) t^n = g(t) + f(t)$$

(In  $(**)$  we used that  $F$  is a field so the addition of elements of  $F$  is commutative.)

**$\cdot$  is associative, commutative, and has an identity 1.**

Closure. For every  $f(t), g(t) \in F[[t]]$ , we trivially have that  $f(t)g(t) \in F[[t]]$ .

Associativity. Let  $f(t), g(t), h(t) \in F[[t]]$ ,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n, \quad g(t) = \sum_{n=0}^{\infty} b_n t^n, \quad h(t) = \sum_{n=0}^{\infty} c_n t^n$$

Then

$$(f(t)g(t))h(t) = \left( \sum_{n=0}^{\infty} d_n t^n \right) h(t) = \sum_{n=0}^{\infty} e_n t^n,$$

where

$$d_n = \sum_{k=0}^n a_k b_{n-k}$$

and

$$e_n = \sum_{l=0}^n d_l c_{n-l} = \sum_{l=0}^n \left( \sum_{k=0}^l a_k b_{l-k} \right) c_{n-l}$$

On the other hand,

$$f(t)(g(t)h(t)) = f(t) \left( \sum_{n=0}^{\infty} x_n t^n \right) = \sum_{n=0}^{\infty} y_n t^n,$$

where

$$\begin{aligned} x_n &= \sum_{k=0}^n b_k c_{n-k} \\ y_n &= \sum_{l=0}^n a_l x_{n-l} = \sum_{l=0}^n x_l a_{n-l} = \sum_{l=0}^n \left( \sum_{k=0}^l b_k c_{l-k} \right) a_{n-l} \end{aligned}$$

Now we rearrange sums a bit:

$$\begin{aligned} y_n &= \sum_{l=0}^n \left( \sum_{k=0}^l b_k c_{l-k} \right) a_{n-l} \\ &= \sum_{l=0}^n \sum_{k=0}^l a_{n-l} b_k c_{l-k} \\ &= \sum_{k=0}^n \sum_{l=0}^k a_l b_{k-l} c_{n-k} \\ &= \sum_{k=0}^n \left( \sum_{l=0}^k a_l b_{k-l} \right) c_{n-k} \\ &= \sum_{l=0}^n \left( \sum_{k=0}^l a_k b_{l-k} \right) c_{n-l} \\ &= e_n \end{aligned}$$

Therefore,

$$(f(t)g(t))h(t) = f(t)(g(t)h(t))$$

Commutativity. Let  $f(t), g(t) \in F[[t]]$ ,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n, \quad g(t) = \sum_{n=0}^{\infty} b_n t^n$$

Then

$$\begin{aligned} f(t)g(t) &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) t^n \\ &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n b_{n-k} a_k \right) t^n \\ &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n b_k a_{n-k} \right) t^n \\ &= g(t)f(t) \end{aligned}$$

Identity. Define

$$id(t) = 1 = 1 + 0 \cdot t + 0 \cdot t^2 + \dots \in F[[t]]$$

Now it is easy to check that it is the multiplicative identity.

#### Distributive law.

Let  $f(t), g(t), h(t) \in F[[t]]$ ,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n, \quad g(t) = \sum_{n=0}^{\infty} b_n t^n, \quad h(t) = \sum_{n=0}^{\infty} c_n t^n$$

Then

$$\begin{aligned} (f(t) + g(t))h(t) &= \left( \sum_{n=0}^{\infty} (a_n + b_n) t^n \right) h(t) \\ &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n (a_k + b_k) c_{n-k} \right) t^n \\ &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n (a_k c_{n-k} + b_k c_{n-k}) \right) t^n \end{aligned}$$

On the other hand

$$\begin{aligned} (f(t)h(t)) + (g(t)h(t)) &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k c_{n-k} \right) t^n + \sum_{n=0}^{\infty} \left( \sum_{k=0}^n b_k c_{n-k} \right) t^n \\ &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n (a_k c_{n-k} + b_k c_{n-k}) \right) t^n \\ &= (f(t) + g(t))h(t) \end{aligned}$$

Therefore, the distributive law holds.

#### Units.

Let  $f(t) \in F[[t]]$ ,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n$$

be a unit. Then there exists some  $g(t) \in F[[t]]$ ,

$$g(t) = \sum_{n=0}^{\infty} b_n t^n$$

such that

$$f(t)g(t) = id(t) = 1 + 0 \cdot t + 0 \cdot t^2 + \dots$$

On the other hand,

$$f(t)g(t) = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) t^n$$

Therefore,

$$\sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) t^n = 1 + 0 \cdot t + 0 \cdot t^2 + \dots$$

Comparing the free coefficients, we get

$$a_0 b_0 = 1$$

From this we get that  $a_0$  has a multiplicative inverse ( $b_0$ ). Since  $F$  is a field, this is equivalent to  $a_0 \neq 0$ .

So, all units must have  $a_0 \neq 0$ . Now suppose that  $f(t) \in F[[t]]$  is a formal power series such that  $a_0 \neq 0$ . We want to prove that it is a unit.

From before, we now that we want to solve the equation

$$\sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) t^n = 1 + 0 \cdot t + 0 \cdot t^2 + \dots$$

for  $b_i$ . We build the sequence  $(b_n)$  inductively.

First of all,  $a_0 b_0 = 1$ , so, since  $a_0 \neq 0$ ,  $b_0 = a_0^{-1}$ .

Now suppose that we know all  $b_n$ , for  $n < m$ , where  $m$  is a positive integer. We want to find  $b_m$ .

To do this, compare the coefficients along  $t^m$ :

$$\sum_{k=0}^m a_k b_{m-k} = 0$$

From this,

$$a_0 b_m = - \sum_{k=1}^m a_k b_{m-k}$$

Therefore, since  $a_0 \neq 0$ ,

$$b_m = -a_0^{-1} \sum_{k=1}^m a_k b_{m-k}$$

Now, if we define

$$g(t) = \sum_{n=0}^{\infty} b_n t^n,$$

we conclude that

$$f(t)g(t) = g(t)f(t) = id(t),$$

so  $f(t)$  is a unit

To conclude,  $f(t) \in F[[t]]$  is a unit if and only if it has a nonzero free coefficient.

## Result

8 of 8

Check properties from Definition 11.1.3.

For the second part,  $f(t) \in F[[t]]$  is a unit if and only if it has a nonzero free coefficient.

## Section 3

1. a

Let  $I$  be some ideal in  $R$ . We check a few properties.

[Subset?](#)

$I \subseteq R^+$  is trivial from the definition of ideals.

[Closure?](#)

Let  $x, y \in I$ . Then  $x + y \in I$  by the definition of an ideal.

[Inverse?](#)

Let  $x \in I$ . Since  $-1 \in R$ , by the definition of an ideal,

$$(-1)x \in I$$

Now,

$$\begin{aligned} x + (-1)x &= 1 \cdot x + (-1)x && \text{(Identity)} \\ &= (1 + (-1))x && \text{(Distributive law)} \\ &= 0x \\ &= 0 \end{aligned}$$

Similarly,  $(-1)x + x = 0$ . Thus,

$$-x = (-1)x,$$

so  $-x \in I$ .

[Conclusion.](#)

Now we conclude that  $I^+$  is a subgroup of  $R^+$ .

---

## Result

2 of 2

Hint: subset and closure is trivial. To check that  $I$  is closed with regard to inverses, prove that  $-x = (-1)x$ .

2. a

Let  $I$  be some nonzero ideal. Then there exists some  $z \in \mathbb{Z}[i]$ ,  $z \neq 0$  such that  $z \in I$ . We know that  $z = a + bi$ , for some  $a, b \in \mathbb{Z}$ . We consider three cases:

1. If  $b = 0$ , then  $z = a$ , so  $z \in \mathbb{Z} \cap I$ , and  $z \neq 0$ , so the statement of the exercise holds.
2. If  $a = 0$ , then  $z = ib$ . Since  $z \neq 0$ , we conclude that  $b \neq 0$ . Since  $I$  is an ideal in  $\mathbb{Z}[i]$ , and  $i \in \mathbb{Z}[i]$ , we conclude that  $iz \in I$ . Furthermore,  $iz = -b \in \mathbb{Z}$ . Thus,  $iz$  is a nonzero integer which is in  $I$ .
3. Let  $a \neq 0$  and  $b \neq 0$ . Since  $I$  is an ideal and  $z \in I$ , we conclude that  $z^2 \in I$ ; that is,

$$(a + bi)^2 = a^2 - b^2 + 2abi \in I$$

Furthermore, since  $-2a \in \mathbb{Z}[i]$ , and  $z \in I$  and  $I$  is an ideal,  $-2az \in I$ ; that is,

$$-2az = -2a(a + bi) = -2a^2 - 2abi \in I$$

Since  $I$  is closed under addition,

$$(a^2 - b^2 + 2abi) + (-2a^2 - 2abi) \in I \implies -a^2 - b^2 \in I$$

Notice that  $-a^2 - b^2 \neq 0$  since  $a^2 > 0$  and  $b^2 > 0$ , so  $-a^2 - b^2 < 0$ . Furthermore, it is an integer. Thus, we have found a nonzero integer in  $I$ .

## Result

HINT: try to find "good enough" elements of  $\mathbb{Z}[i]$  to multiply with nonzero elements of  $I$ .

## 3. a

In all cases, we denote the defined homomorphism as  $\varphi$ .

### (a)

We need to find all polynomials  $f(x, y) \in \mathbb{R}[x, y]$  such that  $\varphi(f(x, y)) = f(0, 0) = 0$ .

The first question is how to write a polynomial in two variables. We will write it as

$$f(x, y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + \dots$$

So, we have a sum

$$f(x, y) = \sum_{n=0}^{\infty} \sum_{i+j=n} a_{ij} x^i y^j$$

Of course, we implicitly assume that only finitely many  $a_{ij}$  are nonzero.

Thus, now

$$\varphi(f(x, y)) = f(0, 0) = a_{00}$$

This means that  $f(0, 0)$  is in  $\ker \varphi$  if and only if  $a_{00} = 0$ , which is if and only if

$$f(x, y) = a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + \dots = xp(x, y) + yq(x, y),$$

where  $p(x, y), q(x, y) \in \mathbb{R}[x, y]$ . Therefore,

$$\ker \varphi = (x, y)$$



(b)

Here we must find all polynomials  $f(x) \in \mathbb{R}[x]$  such that  $\varphi(f(x)) = f(2+i) = 0$ .

This means that  $f(x)$  is in the kernel of  $\varphi$  if and only if  $2+i$  is a root of  $f(x)$ . Since  $f(x)$  has real coefficients, by the *complex conjugate root theorem* we know that  $2-i$  is also a root of  $f(x)$ . This means that  $(x - (2+i))$  and  $(x - (2-i))$  divide  $f(x)$ ; hence,

$$(x - (2+i))(x - (2-i)) = ((x-2) - i)((x-2) + i) = (x-2)^2 - i^2 = x^2 - 4x + 5$$

also divides  $f(x)$ . Since  $f(x)$  and  $x^2 - 4x + 5$  are in  $\mathbb{R}[x]$ , this means precisely that there exists a polynomial  $p(x) \in \mathbb{R}[x]$  such that

$$f(x) = (x^2 - 4x + 5)p(x)$$

Therefore,  $f(x)$  is in the kernel of  $\varphi$  if and only if

$$f(x) = (x^2 - 4x + 5)p(x)$$

for some polynomial  $p(x) \in \mathbb{R}[x]$ . From this we conclude that

$$\ker \varphi = (x^2 - 4x + 5)$$

(c)

Here we must find all polynomials  $f(x) \in \mathbb{Z}[x]$  such that  $\varphi(f(x)) = f(1+\sqrt{2}) = 0$ .

This means that  $f(x)$  is in the kernel of  $\varphi$  if and only if  $1+\sqrt{2}$  is a root of  $f(x)$ . Since  $f(x)$  has integer coefficients, by the *irrational conjugate root theorem* we know that  $1-\sqrt{2}$  is also a root of  $f(x)$ . This means that  $(x - (1+\sqrt{2}))$  and  $(x - (1-\sqrt{2}))$  divide  $f(x)$ ; hence,

$$(x - (1+\sqrt{2}))(x - (1-\sqrt{2})) = ((x-1) - \sqrt{2})((x-1) + \sqrt{2}) = (x-1)^2 - 2 = x^2 - 2x - 1$$

also divides  $f(x)$ . Since  $f(x)$  and  $x^2 - 2x - 1$  are in  $\mathbb{Z}[x]$ , this means precisely that there exists a polynomial  $p(x) \in \mathbb{Z}[x]$  such that

$$f(x) = (x^2 - 2x - 1)p(x)$$

Therefore,  $f(x)$  is in the kernel of  $\varphi$  if and only if

$$f(x) = (x^2 - 2x - 1)p(x)$$

for some polynomial  $p(x) \in \mathbb{Z}[x]$ . From this we conclude that

$$\ker \varphi = (x^2 - 2x - 1)$$

(d)

Here we must find all polynomials  $f(x) \in \mathbb{Z}[x]$  such that  $\varphi(f(x)) = f(\sqrt{2} + \sqrt{3}) = 0$ .

This means that  $f(x)$  is in the kernel of  $\varphi$  if and only if  $\sqrt{2} + \sqrt{3}$  is a root of  $f(x)$ . As in (b) and (c), we want to find all other similar roots of  $f(x)$ .

To lighten notation a bit,

$$f(x) = a_n x^n + \dots + a_1 x + a_0$$

For  $t = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ , with  $a, b, c, d \in \mathbb{Z}$ , define

$$\bar{t} = a + b\sqrt{2} - c\sqrt{3} - d\sqrt{6}$$

(so, we imitate the conjugation of complex numbers somewhat)

We easily check that it satisfies the following properties:

$$\overline{t_1 + t_2} = \bar{t}_1 + \bar{t}_2$$

$$\overline{t_1 t_2} = \bar{t}_1 \cdot \bar{t}_2$$

$$\overline{t_1^n} = \bar{t}_1^n,$$

where  $t_1 = a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}$ ,  $t_2 = a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}$ ,  $a_i, b_i, c_i, d_i \in \mathbb{Z}$  (notice that we need "+d\sqrt{6}" because of the product).

Also, we can now confirm that  $\sqrt{2} - \sqrt{3}$  is a root of  $f(x)$ !

$$\begin{aligned} f(\sqrt{2} - \sqrt{3}) &= \overline{f(\sqrt{2} + \sqrt{3})} \\ &= \overline{a_n \sqrt{2} + \sqrt{3}^n + \dots + a_1 \sqrt{2} + \sqrt{3} + a_0} \\ &= \overline{a_n (\sqrt{2} + \sqrt{3})^n + \dots + a_1 \sqrt{2} + \sqrt{3} + a_0} \\ &= \overline{a_n (\sqrt{2} + \sqrt{3})^n + a_1 (\sqrt{2} + \sqrt{3}) + a_0} \\ &= \overline{f(\sqrt{2} + \sqrt{3})} \\ &= \bar{0} \\ &= 0 \end{aligned}$$

Therefore,  $\sqrt{2} - \sqrt{3}$  is a root of  $f(x)$ .

Similarly, by defining

$$\bar{t} = a - b\sqrt{2} + c\sqrt{3} - d\sqrt{6}$$

we prove that  $-\sqrt{2} + \sqrt{3}$  is a root of  $f(x)$ .

Similarly, by defining

$$\bar{t} = a - b\sqrt{2} - c\sqrt{3} + d\sqrt{6}$$

we prove that  $-\sqrt{2} - \sqrt{3}$  is a root of  $f(x)$ .

Therefore, all of the following polynomials divide  $f(x)$ :

$$(x - (\sqrt{2} + \sqrt{3})), \quad (x - (\sqrt{2} - \sqrt{3})), \quad (x - (-\sqrt{2} + \sqrt{3})), \quad (x - (-\sqrt{2} - \sqrt{3}))$$

Therefore,

$$(x - (\sqrt{2} + \sqrt{3}))(x - (\sqrt{2} - \sqrt{3}))(x - (-\sqrt{2} + \sqrt{3}))(x - (-\sqrt{2} - \sqrt{3}))$$

divides  $f(x)$ . Furthermore,

$$\begin{aligned} & (x - (\sqrt{2} + \sqrt{3}))(x - (\sqrt{2} - \sqrt{3}))(x - (-\sqrt{2} + \sqrt{3}))(x - (-\sqrt{2} - \sqrt{3})) \\ &= ((x - \sqrt{2}) - \sqrt{3})((x - \sqrt{2}) + \sqrt{3})((x + \sqrt{2}) - \sqrt{3})((x + \sqrt{2}) + \sqrt{3}) \\ &= ((x - \sqrt{2})^2 - 3)((x + \sqrt{2})^2 - 3) \\ &= (x^2 - 2\sqrt{2}x - 1)(x^2 + 2\sqrt{2}x - 1) \\ &= ((x^2 - 1) - 2\sqrt{2}x)((x^2 - 1) + 2\sqrt{2}x) \\ &= (x^2 - 1)^2 - 8x^2 \\ &= x^4 - 10x^2 + 1 \end{aligned}$$

Therefore, there exists some polynomial  $p(x) \in \mathbb{Z}[x]$  such that

$$f(x) = (x^4 - 10x^2 + 1)p(x)$$

and

$$\ker \varphi = (x^4 - 10x^2 + 1)$$

(e)

Let  $f(x, y, z) \in \mathbb{C}[x, y, z]$ . By using the division theorem,

$$f(x, y, z) = (x^3 - z)q_1(x, y, z) + (x^2 - y)q_2(x, y) + p(x) \quad (1)$$

In the first division we "eliminate" the variable  $z$  since  $x^3 - z$  is of degree 1 in  $z$ , so the remainder of the division will be constant (of degree 0) in  $z$ . In the second division we repeat the process for  $y$ .

Now,  $f(x, y, z)$  is in kernel of  $\varphi$  if and only if

$$\varphi(f(x, y, z)) = 0$$

which is if and only if

$$f(t, t^2, t^3) = 0$$

Using (1), this holds if and only if

$$p(t) = 0$$

Therefore, using (1) again,  $f(x, y, z) \in \ker \varphi$  if any only if

$$f(x, y, z) = (x^3 - z)q_1(x, y, z) + (x^2 - y)q_2(x, y)$$

Thus,

$$\ker \varphi = (x^3 - z, x^2 - y)$$

## Result

$$\ker \varphi =$$

(a)  $(x, y)$

(b)  $(x^2 - 4x + 5)$

(c)  $(x^2 - 2x - 1)$

(d)  $x^4 - 10x^2 + 1$

(e)  $(x^3 - z, x^2 - y)$

4. a

We will first find a relation between such  $x$  and  $y$ . Firstly,

$$x = t + 1 \implies t = x - 1$$

Therefore,

$$y = t^3 + 1 = (x - 1)^3 + 1 = x^3 - 3x^2 + 3x$$

Now let  $f(x, y) \in \mathbb{C}[x, y]$ . Since  $x^3 - 3x^2 + 3x - y$  can be considered a polynomial of first degree in  $y$ , we can divide  $f(x, y)$  by it to get

$$f(x, y) = (x^3 - 3x^2 + 3x - y)q(x, y) + r(x)$$

(the remainder will be of degree 0 in  $y$ ; thus, it will be a polynomial only of variable  $x$ ).

Now we find the kernel. Notice that  $f(x, y)$  is in kernel if and only if

$$f(t + 1, t^3 - 1) = 0,$$

which is if and only if

$$p(t + 1) = 0$$

Now let

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

Then

$$0 = p(t + 1) = a_n (t + 1)^n + a_{n-1} (t + 1)^{n-1} + \dots + a_1 (t + 1) + a_0 \quad (1)$$

Since

$$(t + 1)^k = \sum_{i=0}^k \binom{k}{i} t^{k-i},$$

from (1), since coefficient along each power must be equal, we get the system of equations

$$\begin{aligned} a_n &= 0 \\ \binom{n}{1} a_n + a_{n-1} &= 0 \\ &\vdots \\ \binom{n}{n-1} a_n + \binom{n-1}{n-2} a_{n-1} + \dots + a_1 &= 0 \\ a_n + a_{n-1} + \dots + a_1 + a_0 &= 0 \end{aligned}$$

From this we see that

$$a_n = a_{n-1} = \dots = a_1 = a_0 = 0$$

Therefore,  $f(x, y)$  is in the kernel if and only if

$$p(x) = 0$$

Therefore,  $f(x, y)$  is in the kernel if and only if

$$f(x, y) = (x^3 - 3x^2 + 3x - y)q(x, y)$$

for some polynomial  $q(x, y) \in \mathbb{C}[x, y]$ . Thus,

$$K = (x^3 - 3x^2 + 3x - y)$$

Now let  $I$  be an ideal which contains  $K$ . Then, for every  $f(x, y) \in I$ , we can write

$$f(x, y) = (x^3 - 3x^2 + 3x - y)q(x, y) + r(x)$$

for some polynomial  $r(x) \in \mathbb{C}[x]$ . This motivates us to define the following set:

$$J = \{r(x) \mid r(x) = f(x, y) - (x^3 - 3x^2 + 3x - y)q(x, y), f(x, y) \in I, q(x, y) \in \mathbb{C}[x, y]\}$$

We will prove that  $J$  is an ideal in  $\mathbb{C}[x]$ . (A good thing to notice now, which we will use later, is that  $r(x) \in I$  since  $I$  is an ideal, and  $f(x, y), x^3 - 3x^2 + 3x - y \in I$ .)

Subset?  $J \subseteq \mathbb{C}[x]$  is trivial.

Closed under addition?

Let  $r_1(x), r_2(x) \in J$ . Then

$$r_1(x) = f_1(x, y) - (x^3 - 3x^2 + 3x - y)q_1(x, y), \quad r_2(x) = f_2(x, y) - (x^3 - 3x^2 + 3x - y)q_2(x, y)$$

for some  $f_1(x, y), f_2(x, y) \in I, q_1(x, y), q_2(x, y) \in \mathbb{C}[x, y]$ . Furthermore,

$$r_1(x) + r_2(x) = (f_1(x, y) + f_2(x, y)) - (x^3 - 3x^2 + 3x - y)(q_1(x, y) + q_2(x, y))$$

Since  $I$  is an ideal,  $f_1(x, y) + f_2(x, y) \in I$ . Therefore,  $r_1(x) + r_2(x) \in I$ .

Closed under subtraction?

Let  $r_1(x), r_2(x) \in J$ . Then

$$r_1(x) = f_1(x, y) - (x^3 - 3x^2 + 3x - y)q_1(x, y), \quad r_2(x) = f_2(x, y) - (x^3 - 3x^2 + 3x - y)q_2(x, y)$$

for some  $f_1(x, y), f_2(x, y) \in I, q_1(x, y), q_2(x, y) \in \mathbb{C}[x, y]$ . Furthermore,

$$r_1(x) - r_2(x) = (f_1(x, y) - f_2(x, y)) - (x^3 - 3x^2 + 3x - y)(q_1(x, y) - q_2(x, y))$$

Since  $I$  is an ideal,  $f_1(x, y) - f_2(x, y) \in I$ . Therefore,  $r_1(x) - r_2(x) \in I$ .



### Closed under multiplication?

Let  $r(x) \in J$ . Then

$$r(x) = f(x, y) - (x^3 - 3x^2 + 3x - y)q(x, y)$$

for some  $f(x, y) \in I$ ,  $q(x, y) \in \mathbb{C}[x, y]$ . Let  $p(x) \in \mathbb{C}[x]$ . Then

$$r(x)p(x) = f(x, y)p(x) - (x^3 - 3x^2 + 3x - y)(q(x, y)p(x))$$

Since  $I$  is an ideal in  $\mathbb{C}[x, y]$ , and  $p(x)$  can be considered a polynomial in  $\mathbb{C}[x, y]$ , we conclude that  $f(x, y)p(x) \in I$ . Moreover, trivially  $q(x, y)p(x) \in \mathbb{C}[x, y]$ . Therefore,  $r(x)p(x) \in J$ .

### Conclusion.

$J$  is an ideal in  $\mathbb{C}[x]$ . By Proposition 11.3.22, we conclude that

$$J = (p(x)),$$

for some polynomial  $p(x) \in \mathbb{C}[x]$ .

Finally, now we see that  $f(x, y) \in I$  if and only if

$$f(x, y) = (x^3 - 3x^2 + 3x - y)q(x, y) + r(x)$$

for some  $q(x, y) \in \mathbb{C}[x, y]$  and  $r(x) \in J$ . Thus,  $r(x) = p(x)g(x)$ , for some  $g(x) \in \mathbb{C}[x]$ . Finally,  $f(x, y) \in I$  if and only if

$$f(x, y) = (x^3 - 3x^2 + 3x - y)q(x, y) + p(x)g(x)$$

Now we prove that  $I$  is generated by  $x^3 - 3x^2 + 3x - y$  and  $p(x)$ . By the above equality,  $I \subseteq (x^3 - 3x^2 + 3x - y, p(x))$  is shown. On the other hand, let  $h(x, y) \in (x^3 - 3x^2 + 3x - y, p(x))$ . Then

$$h(x, y) = (x^3 - 3x^2 + 3x - y)q(x, y) + p(x)g(x, y)$$

for some  $q(x, y), g(x, y) \in \mathbb{C}$ . Now we divide  $g(x, y)$  by  $(x^3 - 3x^2 + 3x - y)$ :

$$g(x, y) = (x^3 - 3x^2 + 3x - y)q_2(x, y) + r(x)$$

Therefore,

$$h(x, y) = (x^3 - 3x^2 + 3x - y)(q(x, y) + p(x)q_2(x, y)) + p(x)r(x)$$

First of all,  $p(x)r(x) \in J$ ; notice that, by definition of  $J$  (using the fact that  $I$  is an ideal), we actually get that  $J \subseteq I$ , so  $p(x)r(x) \in I$ . Furthermore,  $x^3 - 3x^2 + 3x - y \in I$ , so, once again, because  $I$  is an ideal,  $(x^3 - 3x^2 + 3x - y)(q(x, y) + p(x)q_2(x, y)) \in I$ . Finally, we now conclude that  $h(x, y) \in I$ .

Thus, we have proven that  $(x^3 - 3x^2 + 3x - y, p(x)) \subseteq I$ , so

$$I = (x^3 - 3x^2 + 3x - y, p(x))$$

### Result

5 of 5

Hint: let  $f(x, y) \in K$ , and divide it by  $x^3 - 3x^2 + 3x - y$  (this polynomial is obtained using the relation between  $x, y$ , and  $t$ ).

5. a

(a)

Let

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad g(x) = \sum_{n=0}^{\infty} b_n x^n$$

(we implicitly assume that only finitely many  $a_n, b_n$  are nonzero).

Then

$$f(x)g(x) = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) x^n$$

Using the formula provided (written using a summation notation),

$$\begin{aligned} f'(x) &= \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n = \sum_{n=0}^{\infty} c_n x^n \\ g'(x) &= \sum_{n=1}^{\infty} n b_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) b_{n+1} x^n = \sum_{n=0}^{\infty} d_n x^n \\ (f(x)g(x))' &= \sum_{n=1}^{\infty} n \left( \sum_{k=0}^n a_k b_{n-k} \right) x^{n-1} \\ &= \sum_{n=0}^{\infty} (n+1) \left( \sum_{k=0}^{n+1} a_k b_{n+1-k} \right) x^n \\ &= \sum_{n=0}^{\infty} e_n x^n \end{aligned}$$

where

$$c_n = (n+1)a_{n+1}, \quad d_n = (n+1)b_{n+1}, \quad e_n = (n+1) \sum_{k=0}^{n+1} a_k b_{n+1-k}$$

Now,

$$\begin{aligned} f'(x)g(x) &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n c_k b_{n-k} \right) x^n = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n (k+1) a_{k+1} b_{n-k} \right) x^n \\ f(x)g'(x) &= \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k d_{n-k} \right) x^n = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k (n-k+1) b_{n-k+1} \right) x^n \end{aligned}$$

Thus,

$$f'(x)g(x) + f(x)g'(x) = \sum_{n=0}^{\infty} y_n x^n,$$

with

$$y_n = \sum_{k=0}^n (k+1)a_{k+1}b_{n-k} + \sum_{k=0}^n a_k(n-k+1)b_{n-k+1}$$

Now notice that, since  $k+1$  is 0 when  $k = -1$ , we can write

$$\sum_{k=0}^n (k+1)a_{k+1}b_{n-k} = \sum_{k=-1}^n (k+1)a_{k+1}b_{n-k} = \sum_{k=0}^{n+1} k a_k b_{n+1-k}$$

Similarly,  $n-k+1$  is 0 when  $k = n+1$ , so

$$\sum_{k=0}^n a_k(n-k+1)b_{n-k+1} = \sum_{k=0}^{n+1} (n-k+1)a_k b_{n+1-k}$$

Finally, now we get that

$$y_n = \sum_{k=0}^{n+1} [k a_k b_{n+1-k} + (n-k+1)a_k b_{n+1-k}] = \sum_{k=0}^{n+1} (n+1)a_k b_{n+1-k} = e_n$$

From this we conclude that

$$f'(x)g(x) + f(x)g'(x) = (f(x)g(x))',$$

as required.

**(b)**

$\alpha$  is a multiple root of  $f$ .

By definition, since  $\alpha$  is a multiple root of  $f(x)$ ,  $(x - \alpha)^2$  divides  $f(x)$ ; that is, there exists some polynomial  $g(x)$  such that

$$f(x) = g(x)(x - \alpha)^2$$

By **(a)**,

$$\begin{aligned} f'(x) &= g'(x)(x - \alpha)^2 + g(x)(x^2 - 2\alpha x + \alpha^2)' \\ &= g'(x)(x - \alpha)^2 + g(x)(2x - 2\alpha) \\ &= (x - \alpha)[g'(x)(x - \alpha) + 2g(x)] \end{aligned}$$

Thus,  $f'(\alpha) = 0$ , so  $\alpha$  is a common root of  $f$  and  $f'$ .

$\alpha$  is a common root of  $f$  and  $f'$ .

This means that

$$f'(\alpha) = f(\alpha) = 0$$

Since  $f(\alpha) = 0$ , we know that  $x - \alpha$  divides  $f(x)$ ; that is, there exists some polynomial  $h(x)$  such that

$$f(x) = h(x)(x - \alpha) \quad (1)$$

From (a),

$$f'(x) = h'(x)(x - \alpha) + h(x)(x - \alpha)' = h'(x)(x - \alpha) + h(x)$$

Now, since  $f'(\alpha) = 0$ ,

$$0 = f'(\alpha) = h'(\alpha)(\alpha - \alpha) + h(\alpha)$$

From this we get that

$$h(\alpha) = 0$$

Thus,  $x - \alpha$  divides  $h(x)$ , so there exists some polynomial  $p(x)$  such that

$$h(x) = p(x)(x - \alpha)$$

Plugging this into (1), we get

$$f(x) = p(x)(x - \alpha)^2$$

Therefore,  $\alpha$  is a multiple root of  $f$ .

## Result

(a) Hint: this is the equality of polynomials; check that corresponding coefficients are equal.

(b) Use (a) and that  $\alpha$  is a root of  $f(x)$  if and only if  $x - \alpha$  divides  $f(x)$ .

## 6. a

Denote this map by  $\varphi$ .

Homomorphism?

The fact that  $\varphi$  is a homomorphism is trivial from its definition.

Injective?

Let  $p(x, y), q(x, y) \in \mathbb{C}[x, y]$ . We will write  $p(x, y)$  in the following way:

$$p(x, y) = a_0(y) + a_1(y)x + \dots + a_n(y)x^n$$

where  $a_i(y) \in \mathbb{C}[y]$  (this form is easily obtained when distributing various powers of  $x$  in  $p(x, y)$ ). Similarly,

$$q(x, y) = b_0(y) + b_1(y)x + \dots + b_m(y)x^m$$

Now suppose that  $\varphi(p(x, y)) = \varphi(q(x, y))$ . Then

$$a_0(y) + a_1(y)(x + f(y)) + \dots + a_n(y)(x + f(y))^n = b_0(y) + b_1(y)(x + f(y)) + \dots + b_m(y)(x + f(y))^m$$

First of all, the degrees of the polynomials (in variable  $x$ ) on the left and the right side must be equal; hence,  $n = m$ . Furthermore, we consider  $p(x, y)$  and  $q(x, y)$  as polynomials in variable  $x$  with coefficients which are polynomials in variable  $y$ . Using the binomial theorem

$$(x + f(y))^k = \sum_{i=0}^k \binom{k}{i} x^{k-i} (f(y))^i$$

Therefore, the polynomial equality, using the fact that corresponding coefficients must be equal, yields the system of equations

$$\begin{aligned}
a_n(y) &= b_n(y) \\
a_{n-1}(y) + \binom{n}{1}a_n(y)f(y) &= b_{n-1}(y) + \binom{n}{1}b_n(y)f(y) \\
&\vdots \\
a_1(y) + \dots + \binom{n}{n-1}a_n(y)f(y)^{n-1} &= b_1(y) + \dots + \binom{n}{n-1}b_n(y)f(y)^{n-1} \\
a_0(y) + a_1(y)f(y) + \dots + a_n(y)f(y)^n &= b_0(y) + b_1(y)f(y) + \dots + b_n(y)f(y)^n
\end{aligned}$$

From this we easily see that

$$a_i(y) = b_i(y), \quad i = 0, 1, \dots, n$$

Therefore,

$$p(x, y) = q(x, y),$$

so  $\varphi$  is injective.

#### Surjective?

Let  $q(x, y) \in \mathbb{C}[x, y]$ .

$$q(x, y) = b_0(y) + b_1(y)x + \dots + b_n(y)x^n$$

We want to find  $p(x, y) \in \mathbb{C}[x, y]$  such that  $\varphi(p(x, y)) = q(x, y)$ . Since  $\varphi(p(x, y))$  and  $p(x, y)$  have the same degree, we conclude that  $p(x, y)$  is of degree  $n$ . Thus,

$$p(x, y) = a_0(y) + a_1(y)x + \dots + a_n(y)x^n$$

and

$$\varphi(p(x, y)) = a_0(y) + a_1(y)(x + f(y)) + \dots + a_n(y)(x + f(y))^n$$

If we want  $\varphi(p(x, y)) = q(x, y)$ , we get the system (the process is the same as when we proved injectivity):

$$\begin{aligned}
a_n(y) &= b_n(y) \\
a_{n-1}(y) + \binom{n}{1}a_n(y)f(y) &= b_{n-1}(y) \\
&\vdots \\
a_1(y) + \dots + \binom{n-1}{n-2}a_{n-1}(y) + \binom{n}{n-1}a_n(y) &= b_1(y) \\
a_0(y) + a_1(y) + \dots + a_{n-1}(y) + a_n(y) &= b_0(y)
\end{aligned}$$

Therefore, we have a solution!

$$\begin{aligned}
 a_n(y) &= b_n(y) \\
 a_{n-1}(y) &= b_{n-1}(y) - \binom{n}{1} a_n(y) f(y) \\
 &\vdots \\
 a_1(y) &= b_1(y) - \binom{2}{1} a_2(y) f(y) \dots - \binom{n-1}{n-2} a_{n-1}(y) f(y)^{n-2} + \binom{n}{n-1} a_n(y) f(y)^{n-1} \\
 a_0(y) &= b_0(y) - a_1(y) f(y) - \dots - a_{n-1}(y) f(y)^{n-1} - a_n(y) f(y)^n
 \end{aligned}$$

Thus, with such defined coefficients  $a_i(y)$ , we have that  $\varphi(p(x, y)) = q(x, y)$ , which proves that  $\varphi$  is surjective.

## Result

4 of 4

Hint: here it is better to write each polynomial  $p(x, y) \in \mathbb{C}[x, y]$  as

$$p(x, y) = a_0(y) + a_1(y)x + \dots + a_n(y)x^n$$

(so, as a polynomial in variable  $x$  with coefficients which are polynomials in variable  $y$ ).

## 7. a

Let  $\varphi$  be an automorphism. Then it is a homomorphism, and it is bijective; denote its inverse by  $\varphi^{-1}$ . We know that  $\varphi^{-1}$  is also an automorphism.

Now let  $p(x) \in \mathbb{Z}[x]$ ,

$$p(x) = a_0 + a_1x + \dots + a_nx^n$$

Since  $\varphi$  is a homomorphism,

$$\begin{aligned}
 \varphi(p(x)) &= \varphi(a_0 + a_1x + \dots + a_nx^n) \\
 &= \varphi(a_0) + \varphi(a_1x) + \dots + \varphi(a_nx^n) \\
 &= \varphi(a_0) + \varphi(a_1)\varphi(x) + \dots + \varphi(a_n)\varphi(x)^n
 \end{aligned}$$

Moreover,  $\varphi(1) = 1$ , so

$$\begin{aligned}
 \varphi(m) &= \varphi(\underbrace{1 + \dots + 1}_{m \text{ times}}) \\
 &= \underbrace{\varphi(1) + \dots + \varphi(1)}_{m \text{ times}} \\
 &= \underbrace{1 + \dots + 1}_{m \text{ times}} \\
 &= m
 \end{aligned}$$



for all  $m \in \mathbb{N}$ . Furthermore,

$$\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0),$$

which, after adding  $-\varphi(0)$  to both sides, yields  $\varphi(0) = 0$ . Therefore,

$$0 = \varphi(0) = \varphi(m + (-m)) = \varphi(m) + \varphi(-m)$$

Therefore,

$$\varphi(-m) = -\varphi(m) = -m$$

Thus, to get back at our polynomial  $p(x)$ , for each  $a_i$  we have that  $\varphi(a_i) = a_i$ , hence

$$\varphi(p(x)) = a_0 + a_1\varphi(x) + \dots + a_n\varphi(x)^n$$

So, the only question which remains is:  $\varphi(x) = ?$

First of all,  $\varphi(x) \in \mathbb{Z}[x]$ ; thus,

$$\varphi(x) = b_0 + b_1x + \dots + b_mx^m$$

However, notice that  $\varphi(p(x))$  is then of degree  $nm$ ; thus, if  $m > 1$ ,  $\varphi$  cannot be an automorphism! For example, we cannot possibly find a polynomial  $p(x)$  such that  $\varphi(p(x)) = x$ , since  $p(x)$  is of degree  $mn$ , which is not 1 for any  $n \in \mathbb{N} \cup \{0\}$ .

Therefore,

$$\varphi(x) = b_0 + b_1x,$$

for some  $b_0, b_1 \in \mathbb{Z}$ .

Similarly,

$$\varphi^{-1}(k) = k$$

for  $k \in \mathbb{Z}$ , and

$$\varphi^{-1}(x) = c_0 + c_1x$$

for some  $c_0, c_1 \in \mathbb{Z}$ .

Moreover,

$$\varphi^{-1}(\varphi(x)) = x$$

and

$$\varphi^{-1}(\varphi(x)) = \varphi^{-1}(b_0 + b_1x) = b_0 + b_1(c_0 + c_1x) = b_0 + b_1c_0 + b_1c_1x$$

Therefore, since

$$b_0 + b_1c_0 + b_1c_1x = x,$$

we conclude that

$$b_1c_1 = 1$$

Since  $b_1, c_1$  are integers, we get that

$$\boxed{b_1 = \pm 1}$$

Now we want to prove that all homomorphisms such that

$$\varphi(x) = b \pm x,$$

for some  $b \in \mathbb{Z}$ , are automorphisms.

### Homomorphism?

This is trivial because we defined  $\varphi$  as a homomorphic extension of the map  $k \mapsto k, k \in \mathbb{Z}$ , and  $x \mapsto b \pm x$ .

### Injective?

Let  $p(x), q(x) \in \mathbb{Z}$ ,

$$p(x) = a_0 + a_1x + \dots + a_nx^n$$

$$q(x) = b_0 + b_1x + \dots + b_mx^m$$

be such that  $\varphi(p(x)) = \varphi(q(x))$ . Then

$$a_0 + a_1(b \pm x) + \dots + a_n(b \pm x)^n = b_0 + b_1(b \pm x) + \dots + b_m(b \pm x)^m$$

Since the degrees of both sides must be equal, we get that  $n = m$ . Furthermore, recall the binomial theorem

$$(b + x)^k = \sum_{i=0}^k \binom{k}{i} b^{k-i} x^i,$$

$$(b - x)^k = \sum_{i=0}^k \binom{k}{i} b^{k-i} (-1)^i x^i$$

Suppose that  $\varphi(x) = b + x$ . Using the binomial theorem and the fact that corresponding coefficients must be equal, we obtain the system of equations

$$\begin{aligned} a_n &= b_n \\ a_{n-1} + \binom{n}{n-1} a_n b &= b_{n-1} + \binom{n}{n-1} b_n b \\ &\vdots \\ a_1 + \binom{2}{1} a_2 b + \dots + \binom{n}{1} b^{n-1} a_n &= b_1 + \binom{2}{1} b_2 b + \dots + \binom{n}{1} b^{n-1} b_n \\ a_0 + a_1 + \dots + a_{n-1} + a_n &= b_0 + b_1 + \dots + b_{n-1} + b_n \end{aligned}$$

From this we see that  $\boxed{a_i = b_i, \quad i = 0, 1, \dots, n}$ .

Similarly, if  $\varphi(x) = b - x$ , then the system is

$$\begin{aligned} a_n &= b_n \\ (-1)^{n-1} a_{n-1} + (-1)^{n-1} \binom{n}{n-1} a_n b &= (-1)^{n-1} b_{n-1} + (-1)^{n-1} \binom{n}{n-1} b_n b \\ &\vdots \\ -a_1 - \binom{2}{1} a_2 b - \dots - \binom{n}{1} b^{n-1} a_n &= -b_1 - \binom{2}{1} b_2 b - \dots - \binom{n}{1} b^{n-1} b_n \\ -a_0 - a_1 - \dots - a_{n-1} - a_n &= -b_0 - b_1 - \dots - b_{n-1} - b_n \end{aligned}$$

which becomes the same system as before!

Therefore,  $p(x) = q(x)$ , and  $\varphi$  is injective.

### Surjective?

Let  $q(x) \in \mathbb{Z}[x]$ ,

$$q(x) = b_0 + b_1x + \dots + b_nx^n$$

We want to find  $p(x) \in \mathbb{Z}[x]$  such that

$$\varphi(p(x)) = q(x)$$

Suppose that  $\varphi(x) = b + x$ . Since the degree of  $\varphi(p(x))$  equals the degree of  $p(x)$ , we get that  $p(x)$  is of degree  $n$ ; thus,

$$p(x) = a_0 + a_1x + \dots + a_nx^n$$

Thus,  $\varphi(p(x)) = q(x)$  becomes

$$a_0 + a_1(b + x) + \dots + a_n(b + x)^n = b_0 + b_1x + \dots + b_nx^n$$

As when proving injectivity, this yields the system

$$\begin{aligned} a_n &= b_n \\ a_{n-1} + \binom{n}{n-1}a_nb &= b_{n-1} \\ &\vdots \\ a_1 + \binom{2}{1}a_2b + \dots + \binom{n}{1}b^{n-1}a_n &= b_1 \\ a_0 + a_1 + \dots + a_{n-1} + a_n &= b_0 \end{aligned}$$

This system has a solution!

$$\begin{aligned} a_n &= b_n \\ a_{n-1} &= b_{n-1} - \binom{n}{n-1}a_nb \\ &\vdots \\ a_1 &= b_1 - \binom{2}{1}a_2b - \dots - \binom{n}{1}b^{n-1}a_n \\ a_0 &= b_0 - a_1 - \dots - a_{n-1} - a_n \end{aligned}$$

Similarly, if  $\varphi(x) = b - x$ , then the system is

$$\begin{aligned} a_n &= b_n \\ (-1)^{n-1}a_{n-1} + (-1)^{n-1}\binom{n}{n-1}a_nb &= b_{n-1} \\ &\vdots \\ -a_1 - \binom{2}{1}a_2b - \dots - \binom{n}{1}b^{n-1}a_n &= b_1 \\ -a_0 - a_1 - \dots - a_{n-1} - a_n &= b_0 \end{aligned}$$

This can also be seen that it has a solution.

Therefore, we can pick  $a_i, i = 0, 1, \dots, n$ , such that

$$\varphi(p(x)) = q(x),$$

so  $\varphi$  is surjective.

## Step 6

6 of 7

### CONCLUSION!

From the beginning, for  $\varphi$  to even have a chance to be an automorphism, it must be a mapping for which  $\varphi(x) = b \pm x$ , for some integer  $b$ ,  $\varphi(1) = 1$ , and we must extend it to be a homomorphism.

On the other hand, all mappings defined by this way are automorphisms.

## Result

7 of 7

All mappings of the form  $\varphi(1) = 1$ , and for some integer  $b$   $\varphi(x) = b \pm x$ , which we then extend on the whole ring  $\mathbb{Z}[x]$ .

## 8. a

Denote this map by  $\varphi$ . We must check that, for every  $x, y \in R$ , we have that

$$\varphi(x + y) = \varphi(x) + \varphi(y), \quad \varphi(xy) = \varphi(x)\varphi(y), \quad \varphi(1) = 1$$

### Multiplicative identity?

Since

$$\varphi(1) = 1^p = 1,$$

$\varphi$  maps the multiplicative identity to itself.

### Multiplication?

Since multiplication on  $R$  is commutative by definition,

$$\varphi(xy) = (xy)^p = \underbrace{(xy)(xy) \cdots (xy)}_{p \text{ times}} = x^p y^p = \varphi(x)\varphi(y)$$

### Addition?

Let  $x, y \in R$ . Then

$$\varphi(x + y) = (x + y)^p$$

By the binomial theorem,

$$(x + y)^p = \sum_{i=0}^p \binom{p}{i} x^{p-i} y^i$$

Furthermore, all numbers  $\binom{p}{i}$  are integers! We have that

$$\binom{p}{i} = \frac{p!}{i!(p-i)!}$$

Let  $i \neq 0, i \neq p$ . Then all numbers which form products  $i!$  and  $(p-i)!$  are less than  $p$ , so none can divide it (because  $p$  is prime). Therefore,

$$\binom{p}{i} = p \cdot k,$$

for some integer  $k > 0$ . Therefore, since  $R$  is of characteristic  $p$ ,

$$\binom{p}{i} x^{p-i} y^i = p k x^{p-i} y^i = k \underbrace{(p x^{p-i} y^i)}_0 = 0$$

For  $i = 0$  or  $i = p$  we have that  $\binom{p}{i} = 1$ . Finally, with all this, we conclude that

$$\varphi(x + y)^p = \sum_{i=0}^p \binom{p}{i} x^{p-i} y^i = x^p + y^p = \varphi(x) + \varphi(y)$$

### Conclusion.

We checked all properties from the definition of a ring homomorphism, thus now we conclude that  $\varphi$  is a homomorphism.

## Result

3 of 3

Hint: Denote this map by  $\varphi$ . The multiplicative identity and multiplication property follow from the definition of  $\varphi$  and ring, respectively. To prove the remaining property, use the binomial theorem.

9. a

(a)

Suppose that  $x^n = 0$ . Then

$$(1 + x)(1 - x + x^2 - \dots + (-1)^{n-1} x^{n-1}) = 1 + (-1)^{n-1} \underbrace{x^n}_0 = 1 + 0 = 1$$

Therefore,

$$(1 + x)^{-1} = 1 - x + x^2 - \dots + (-1)^{n-1} x^{n-1}$$

so it is a unit.

(b)

We first prove that

$$(1 + a)^p = 1 + a^p$$

By the binomial theorem,

$$(x + y)^p = \sum_{i=0}^p \binom{p}{i} x^{p-i} y^i$$

Furthermore, all numbers  $\binom{p}{i}$  are integers! We have that

$$\binom{p}{i} = \frac{p!}{i!(p-i)!}$$

Let  $i \neq 0, i \neq p$ . Then all numbers which form products  $i!$  and  $(p-i)!$  are less than  $p$ , so none can divide it (because  $p$  is prime). Therefore,

$$\binom{p}{i} = p \cdot k,$$

for some integer  $k > 0$ . Therefore, since  $R$  is of characteristic  $p$ ,

$$\binom{p}{i} x^{p-i} y^i = p k x^{p-i} y^i = k \underbrace{(p x^{p-i} y^i)}_0 = 0$$

For  $i = 0$  or  $i = p$  we have that  $\binom{p}{i} = 1$ . Finally,

$$(1 + a)^p = \sum_{i=0}^p \binom{p}{i} a^i = 1 + a^p$$

Inductively, for all positive integers  $m$ ,

$$(1 + a)^{p^m} = ((1 + a)^p)^{p^{m-1}} = (1 + a^p)^{p^{m-1}} = \dots = 1 + a^{p^m}$$

Now let  $n$  be such positive integer such that  $a^n = 0$ . Then, multiplying this equality by  $a$ , we get  $a^{n+1} = 0$ .

Inductively, we see that  $a^k = 0$ , for all  $k \geq n$ .

Now we observe numbers of the form  $p^m$ , where  $m$  is a positive integer. For a sufficiently large  $m$ , we have that  $p^m \geq n$ ; that is,  $a^{p^m} = 0$ . From before,

$$(1 + a)^{p^m} = 1 + a^{p^m} = 1 + 0 = 1$$

Therefore,  $1 + a$  is unipotent.

## Result

3 of 3

(a) Show that  $(1 + x)^{-1} = (1 - x + x^2 - \dots + (-1)^{n-1} x^{n-1})$

(b) Hint: prove that  $(1 + a)^{p^m} = 1 + a^{p^m}$ .



First notice that  $(t^m)$ , for any nonnegative (possibly 0) integer  $m$  is an ideal in  $F[[t]]$ . We will prove that those are all nonzero ideals in  $F[[t]]$ .

Let  $I$  be some nonzero ideal of  $F[[t]]$ . Suppose that all  $f(t)$  in  $F[[t]]$  are of the form

$$f(t) = a_m t^m + a_{m+1} t^{m+1} + \dots,$$

for some nonnegative integer  $m$ , and that there is some element of  $I$  such that  $a_m \neq 0$  (such exists because we assumed that  $I$  is a nonzero ideal). First of all, we can write

$$f(t) = t^m (a_m + a_{m+1} t + \dots)$$

which proves that  $f(t) \in (t^m)$ . Therefore,

$$I \subseteq (t^m)$$

Now let  $g(t) \in I$ ,

$$g(t) = b_m t^m + b_{m+1} t^{m+1} + \dots,$$

with  $b_m \neq 0$ . We write

$$g(t) = t^m (b_m + b_{m+1} t + \dots)$$

Since  $b_m \neq 0$ , the power series

$$b_m + b_{m+1} t + \dots$$

is a unit in  $F[[t]]$ ! (Check Exercise 2.2.)

Thus, there exists some  $u(t)$  such that

$$(b_m + b_{m+1} t + \dots) u(t) = 1$$

This means that

$$t^m = g(t) u(t)$$

Since  $I$  is an ideal and  $g(t) \in I$ , we conclude that  $t^m \in I$ . Furthermore, let  $h(t) \in (t^m)$ . Then

$$h(t) = t^m h_2(t),$$

for some  $h_2(t) \in F[[t]]$ . Finally, since  $t^m \in I$  and  $I$  is an ideal, we conclude that  $h(t) \in I$ . Finally,

$$(t^m) \subseteq I$$

Now we conclude that  $I = (t^m)$ . Since  $I$  was an arbitrary taken nonzero ideal, all nonzero ideals in  $F[[t]]$  are of this form.

## Result

Zero ideal and principal ideals  $(t^m)$ , where  $m$  is a nonnegative integer (possibly 0).

11. a

This statement is false.

Take the ring  $\mathbb{Z}[x]$ . Consider the ideal

$$I = (2)$$

First of all,

$$I = \{2p(x) \mid p(x) \in \mathbb{Z}[x]\}$$

Clearly, the smallest degree among nonzero polynomials is 0. We will prove that  $1 \notin I$ , which is the only monic polynomial of degree 0.

Suppose that

$$1 \in I$$

Then there exists some polynomial  $p(x) \in \mathbb{Z}[x]$  such that

$$1 = 2p(x)$$

Since  $2p(x)$  is of degree  $\deg p$ , where  $\deg p$  is the degree of  $p$ , we conclude that we must have  $\deg p = 0$ . Therefore,  $p(x) = c$ , for some  $c \in \mathbb{Z}$ ,  $c \neq 0$ . Therefore,

$$2c = 1$$

which is absurd.

Therefore,  $I$  does not contain a monic polynomial of degree 0.

## Result

The statement is false. Take  $\mathbb{Z}[x]$  and the ideal  $(2)$ .

12. a

### Nonempty subset?

Since  $I \neq \emptyset$  and  $J \neq \emptyset$  (they are also ideals of  $R$ ), we conclude that  $I + J \neq \emptyset$ . This is because we have some  $x \in I$  and  $y \in J$ , so  $x + y \in I + J$ .

The fact that  $I + J \subseteq R$  is trivial;  $I \subseteq R$  and  $J \subseteq R$  means that for each  $x \in I$  we have that  $x \in R$  and for each  $y \in J$  we have that  $y \in R$ . Therefore, for every  $x + y \in I + J$  we have that  $x + y \in R$  since  $R$  is closed under addition.

### Closed under addition?

Let  $z_1, z_2 \in I + J$ . Then there exist some  $x_1, x_2 \in I$ ,  $y_1, y_2 \in J$  such that

$$z_1 = x_1 + y_1$$

$$z_2 = x_2 + y_2$$

Now,

$$z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)$$

since addition in the ring is commutative. Furthermore, since  $I, J$  are ideals,  $x_1 + x_2 \in I$  and  $y_1 + y_2 \in J$ . Therefore,  $z_1 + z_2 \in I + J$ .

#### Closed under multiplication?

Let  $r \in R$ ,  $s \in I + J$ . Then there exist some  $x \in I$ ,  $y \in J$  such that  $s = x + y$ . Furthermore, because of the distributive law,

$$rs = r(x + y) = rx + ry$$

Since  $I$  is an ideal in  $R$  and  $x \in I$ ,  $rx \in I$ . Similarly,  $ry \in J$ . Thus,  $rs \in I + J$ , as required.

#### **Result**

2 of 2

Check the properties from the Definition 11.3.13.

### 13. a

#### Intersection.

We check the properties from the definition of an ideal.

#### Nonempty subset?

Since  $I \subseteq R$  and  $J \subseteq R$ , we know that  $I \cap J \subseteq R$ . Moreover, it is easy to see that each ideal contains 0, so  $0 \in I$  and  $0 \in J$ , which in turn means that  $0 \in I \cap J$  and  $I \cap J \neq \emptyset$ .

#### Closed under addition?

Let  $x, y \in I \cap J$ . Then  $x, y \in I$ , so  $x + y \in I$  since  $I$  is an ideal. Similarly, since  $x, y \in J$ , and  $J$  is an ideal,  $x + y \in J$ . Finally, this means that  $x + y \in I \cap J$ .

#### Closed under multiplication?

Let  $x \in I \cap J$ ,  $r \in R$ . Since  $x \in I$ , and  $I$  is an ideal,  $rx \in I$ . Similarly,  $x \in J$ , and  $J$  is an ideal, so  $rx \in J$ . Finally, we conclude that  $rx \in I \cap J$ .

#### Conclusion.

We conclude that  $I \cap J$  is an ideal in  $R$ .

#### Set of products is not an ideal.

Consider the ring  $\mathbb{R}[x, y]$ , and ideals  $I = (x, y)$ ,  $J = (x, y)$ . Then

$$I = J = \{xp(x, y) + yq(x, y) \mid p(x, y), q(x, y) \in \mathbb{R}[x, y]\}$$

Denote by  $K$  the set of products:

$$K = \{ab \mid a \in I, b \in J\}$$

Then it is clear that  $x^2 \in K$ ,  $y^2 \in K$ , but  $x^2 + y^2 \notin K$ , so  $K$  is not closed under addition, and hence it is not an ideal.

### Product ideal.

To prove that  $IJ$  is an ideal, we check the properties from the definition of an ideal.

### Nonempty subset?

$IJ \subseteq R$  is trivial from the definition, since all products  $xy \in R$ , and  $R$  is closed under addition. Since  $I, J$  are ideals, they are nonempty. Thus there exist some  $x \in I, y \in J$ . Now  $xy \in IJ$ , so  $IJ \neq \emptyset$ .

### Closed under addition?

Let  $a, b \in IJ$ . Then  $x = \sum_{i=1}^n x_i y_i, y = \sum_{i=n+1}^{n+m} x_i y_i$ , for some  $x_i \in I, y_i \in J$ . Thus,  $x + y = \sum_{i=1}^{n+m} x_i y_i$ . Finally, this means that  $x + y \in IJ$ .

### Closed under multiplication?

Let  $x \in IJ, r \in R$ . Then  $x = \sum_{i=1}^n x_i y_i$ , for some  $x_i \in I, y_i \in J$ . Using the distributive law and the associativity of multiplication,

$$rx = r \sum_{i=1}^n x_i y_i = \sum_{i=1}^n (rx_i) y_i$$

Since  $I$  is an ideal, we conclude that  $rx_i \in I$ . Finally, we conclude that  $rx \in IJ$ .

### Conclusion.

We conclude that  $IJ$  is an ideal in  $R$ .

### Relation between $I \cap J$ and $IJ$ .

We will prove that  $IJ \subseteq I \cap J$ .

Let  $a \in IJ$ . Then

$$a = \sum_{i=1}^n x_i y_i,$$

for some  $x_i \in I, y_i \in J$ . Since  $J$  is an ideal, we conclude that  $x_i y_i \in J$ , for  $i = 1, 2, \dots, n$ . Thus,

$$a = \sum_{i=1}^n x_i y_i \in J$$

since  $J$  is closed under addition.

On the other hand, since  $I$  is an ideal, we conclude that  $x_i y_i \in I$ , for  $i = 1, 2, \dots, n$ . Thus,

$$a = \sum_{i=1}^n x_i y_i \in I$$

since  $I$  is closed under addition.

Finally, this means that  $a \in I \cap J$ . Since  $a$  was arbitrarily picked, we conclude that

$$\boxed{IJ \subseteq I \cap J}$$

The equality (and the other inclusion with it) need not hold. Consider  $R = \mathbb{Z}$ ,  $I = J = 2\mathbb{Z}$  (so, the sets of even integers). Then  $I \cap J = 2\mathbb{Z}$ .

On the other hand, let  $a \in IJ$ . Then

$$a = \sum_{i=1}^n x_i y_i$$

Since  $x_i y_i$  will be divisible by 4,  $a$  will also be divisible by 4. Thus,  $a \in 4\mathbb{Z}$ . Therefore,  $IJ \subseteq 4\mathbb{Z}$  (the equality also holds but we do not need to show it here).

Since  $2 \notin 4\mathbb{Z}$ , we have  $4\mathbb{Z} \subset 2\mathbb{Z}$  (a proper subset). Thus,

$$IJ \subseteq 4\mathbb{Z} \subset 2\mathbb{Z} = I \cap J,$$

which shows that  $IJ \neq I \cap J$ .

## Result

4 of 4

To show that  $I + J$  and  $IJ$  are ideals, check properties from the definition of an ideal.

To show that the set of products of elements of  $I$  and  $J$  is not ideal, consider the ring  $\mathbb{R}[x, y]$  with  $I = J = (x, y)$ .

For the final question, show that  $IJ \subseteq I \cap J$ , and that the equality need not hold.

## Section 4

1. a

We will first find the kernel of this homomorphism, which we will denote by  $\varphi$ .

First of all, clearly

$$\varphi(m) = m,$$

for all  $m \in \mathbb{Z}$ , because  $\varphi$  is closed under addition and multiplication by  $-1$ .

Let  $f(x) \in \ker \varphi$ . Since  $\varphi(f(x)) = 0$ , from the definition of  $\varphi$  we get that  $f(1) = 0$ . Thus, 1 is a root of  $f(x)$ ! Therefore,  $x - 1$  divides  $f(x)$ ; that is,

$$f(x) = (x - 1)g(x)$$

for some polynomial  $g(x) \in \mathbb{Z}[x]$ . This means that

$$\ker \varphi \subseteq (x - 1)$$

To prove the other inclusion, let  $h(x) \in (x - 1)$ . Then there exists some polynomial  $p(x)$  such that  $h(x) = p(x)(x - 1)$ . Now it is clear that  $h(1) = 0$ , so  $\varphi(h(x)) = 0$ , and  $h(x) \in \ker \varphi$ .

Finally,

$$\ker \varphi = (x - 1)$$

Now denote  $K = \ker \varphi$ . Using the Correspondence Theorem, we can see what ideals  $I$  which contain  $K$  are. Let  $I$  be some ideal of  $\mathbb{Z}[x]$  which contains  $K$ . By the Correspondence Theorem, there exists some ideal  $J$  in  $\mathbb{Z}$  such that

$$\varphi^{-1}(J) = I$$

Since  $J$  is an ideal in  $\mathbb{Z}$ , we know that  $J$  is a principle ideal; that is, there exists some  $m \in \mathbb{Z}$  such that  $J = (m)$ .

Now,

$$J = \{km \mid k \in \mathbb{Z}\},$$

so

$$\varphi^{-1}(J) = \{f(x) \mid \varphi(f(x)) = km, k \in \mathbb{Z}\}$$

Now let

$$f(x) = \sum_{i=0}^n a_i x^i$$

be some polynomial in  $\mathbb{Z}[x]$ . Then

$$\varphi(f(x)) = \sum_{i=0}^n a_i$$

Therefore,  $\varphi$  maps  $f(x)$  to the sum of its coefficients!

Therefore,

$$\varphi(f(x)) = n$$

if and only if the sum of coefficients of  $f(x)$  is equal to  $n$ . Now we conclude that

$$\{p(x) \mid \varphi(p(x)) = km, k \in \mathbb{Z}\}$$

is the set of all polynomials whose sum of coefficients is a multiple of  $m$ .

Therefore,  $I$  is a set of polynomials whose sum of coefficients is a multiple of  $m$ . Since  $I$  was an arbitrarily taken ideal which contains  $K$ , all ideals which contain  $K$  are of this form (for some, maybe other, integer  $m$ ).



**Result**

3 of 3

Show that the kernel of this map is  $K = (x - 1)$ . Then we can show that all ideals which contain  $K$  are the sets of polynomials with the sum of coefficients equal to some multiple of  $m$ , where  $m$  is some integer.

2. a

Consider the homomorphism  $\varphi : \mathbb{Z}[x] \rightarrow \mathbb{Z}[i]$  defined by  $x \mapsto i$ . First we prove that  $\ker \varphi = (x^2 + 1)$ .

Let  $f(x) \in \mathbb{Z}[x]$  be such that  $f(x) \in \ker \varphi$ . Then  $\varphi(f(x)) = 0$ . This is equivalent to  $f(i) = 0$ . Thus,  $i$  is a root of  $f(x)$ . By the complex conjugate theorem, we know that  $-i$  is also a root of  $f(x)$ . Thus, the polynomial

$$(x + i)(x - i) = x^2 + 1$$

divides  $f(x)$ . This means that

$$f(x) = g(x)(x^2 + 1)$$

for some  $g(x) \in \mathbb{Z}[x]$ . Thus,

$$\ker(\varphi) \subseteq (x^2 + 1)$$

Now let  $h(x) \in (x^2 + 1)$ . Then there exists some  $h_2(x)$  such that  $h(x) = h_2(x)(x^2 + 1)$ . Thus,  $h(i) = 0$ , which means that  $\varphi(h(x)) = 0$ , and  $h(x) \in \ker(\varphi)$ . Thus,

$$(x^2 + 1) \subseteq \ker(\varphi)$$

Finally, we conclude that

$$\ker(\varphi) = (x^2 + 1)$$

which we wanted to show.

Now, by the Correspondence Theorem, the ideals which contain  $(x^2 + 1)$  correspond to ideals in  $\mathbb{Z}[i]$ .

**Result**

They correspond to ideals in  $\mathbb{Z}[i]$ .

3. a

(a)

Let  $I = (x^2 - 3, 2x + 4)$ . Then  $f(x) \in I$  if and only if there exist polynomials  $p(x) \in \mathbb{Z}[x]$ ,  $q(x) \in \mathbb{Z}[x]$  such that

$$f(x) = p(x)(x^2 - 3) + q(x)(2x + 4)$$

Now if we take  $p(x) = 2$ ,  $q(x) = 2 - x$ , we get

$$2(x^2 - 3) + (2 - x)(2x + 4) = 2x^2 - 6 + 4x + 8 - 2x^2 - 4x = 2$$

Thus,  $2 \in I$ .

Now  $4 = 2 \cdot 2 \in I$ , so

$$(x^2 - 3) + 4 = x^2 + 1 \in I$$

We will prove that

$$I = (x^2 + 1, 2)$$

First of all, clearly

$$(x^2 + 1, 2) \subseteq I,$$

since  $(x^2 + 1), 2 \in I$ , and  $(x^2 + 1, 2)$  is the smallest ideal which contains those two elements.

On the other hand,

$$x^2 + 1 + (-2)2 = x^2 - 3,$$

so  $x^2 - 3 \in (x^2 + 1, 2)$ .

Also,

$$(x + 2)2 = 2x + 4 \in (x^2 + 1, 2)$$

Using the same argument as before,

$$I = (x^2 - 3, 2x + 4) \subseteq (x^2 + 1, 2)$$

Therefore,

$$I = (x^2 + 1, 2)$$

So, what we will identify is

$$\mathbb{Z}[x]/(x^2 + 1, 2)$$

From Example 11.4.5, we know that

$$\mathbb{Z}[x]/(x^2 + 1) \approx \mathbb{Z}[i]$$

Since the image of 2 by homomorphism  $x \rightsquigarrow i$  is 2, we get that

$$\mathbb{Z}[x]/(x^2 + 1, 2) \approx \mathbb{Z}[i]/(2)$$

Now it is easy to see that

$$\mathbb{Z}[i]/(2) \approx (\mathbb{Z}/2\mathbb{Z})[i]$$

Truly, define a homomorphism

$$f(a + bi) = \bar{a} + \bar{b}i,$$

where  $a, b \in \mathbb{Z}, \bar{a}, \bar{b} \in \mathbb{Z}/2\mathbb{Z}$ . This is a surjective homomorphism with kernel  $(2)$ , because both  $a$  and  $b$  must be even for  $a + bi$  to be in the kernel. Thus, by the First Isomorphism Theorem,

$$\mathbb{Z}[i]/(2) \approx (\mathbb{Z}/2\mathbb{Z})[i]$$

Finally, we now get that

$$\boxed{\mathbb{Z}/(x^2 - 3, 2x + 4) \approx (\mathbb{Z}/2\mathbb{Z})[i]}$$

## (b)

This is similar to Example 11.4.5. First of all,

$$\mathbb{Z}/(x^2 + 1) \approx \mathbb{Z}[i]$$

Since the image of  $x + 2$  by the function  $x \rightsquigarrow i$  is  $2 + i$ , we conclude that

$$\mathbb{Z}/(x^2 + 1, x + 2) \approx \mathbb{Z}[i]/(2 + i)$$

Now the homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z} \ x \rightsquigarrow -2$  has the kernel  $(x + 2)$ . On the other hand, this homomorphism is clearly surjective, so

$$\mathbb{Z}[x]/(x + 2) \approx \mathbb{Z}$$

by the First Isomorphism Theorem.

The residue of  $x^2 + 1$  is 5, so we now kill 5 in  $\mathbb{Z}$ , meaning that

$$\boxed{\mathbb{Z}[i]/(i + 2) \approx \mathbb{Z}[x]/(x^2 + 1, x + 2) \approx \mathbb{Z}/5\mathbb{Z}}$$

(c)

Let  $I = (6, 2x - 1)$ . Since

$$3 = 6 \cdot x + (2x - 1)(-3),$$

we conclude that  $3 \in I$ . Furthermore,

$$(2x - 1) \cdot (-1) + 3 \cdot x = x + 1 \in I,$$

since  $I$  is an ideal, so  $3x$ ,  $-(2x - 1)$ , and their sum is in  $I$ .

Now we prove that

$$I = (3, x + 1)$$

First of all, since  $3, (x + 1) \in I$ , and  $(3, x + 1)$  is the smallest ideal which contains these elements, we get

$$(3, x + 1) \in I$$

Similarly,

$$6 = 3 \cdot 2 \in (3, x + 1)$$

and

$$2x - 1 = (x + 1) \cdot 2 + 3 \cdot (-1) \in (3, x + 1)$$

Thus,

$$I = (6, 2x - 1) \subseteq (3, x + 1)$$

and

$$I = (3, x + 1)$$

Thus,

$$\mathbb{Z}[x](6, 2x - 1) = \mathbb{Z}[x](3, x + 1)$$

Now we consider a homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z}$  which sends  $x \mapsto -1$ . Its kernel is  $(x + 1)$  and it is clearly surjective. Thus, by the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(x + 1) \approx \mathbb{Z}$$

The image of 3 under this homomorphism is 3, so we kill 3 in  $\mathbb{Z}$ , from which we get

$$\mathbb{Z}[x]/(x + 1, 3) \approx \mathbb{Z}/3\mathbb{Z}$$

Finally,

$$\boxed{\mathbb{Z}[x]/(6, 2x - 1) \approx \mathbb{Z}/3\mathbb{Z}}$$

(d)

Let  $I = (2x^2 - 4, 4x - 5)$ . First of all,

$$(2x^2 - 4) \cdot 2 + (4x - 5)(-x - 1) = 4x^2 - 8 - 4x^2 + x + 5 = x - 3$$

So,  $x - 3 \in I$ . Now,

$$(x - 3) \cdot (-4) + (4x - 5) = 7,$$

so  $7 \in I$ , since  $I$  is an ideal, so it is closed under addition and ring multiplication.

Now we will prove that  $I = (x - 3, 7)$ . Since  $(x - 3), 7 \in I$ , and  $(x - 3, 7)$  is the smallest ideal which contains these elements, we conclude that

$$(x - 3, 7) \subseteq I$$

On the other hand,

$$2x^2 - 4 = (x - 3)(-5x - 1) + 7(x^2 - 2x - 1)$$

and

$$4x - 5 = (x - 3) \cdot 4 + 7$$

so

$$I = (2x^2 - 4, 4x - 5) \subseteq (x - 3, 7)$$

Thus,

$$I = (x - 3, 7)$$

and

$$\mathbb{Z}[x]/(2x^2 - 4, 4x - 5) = \mathbb{Z}[x]/(x - 3, 7)$$

We now consider the homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z}$  which sends  $x \rightsquigarrow 3$ . Its kernel is  $(x - 3)$ , and it is clearly surjective. Thus, by the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(x - 3) \approx \mathbb{Z}$$

Moreover, the residue of 7 is 7, so we kill 7 in  $\mathbb{Z}$ , obtaining

$$\mathbb{Z}[x]/(x - 3, 7) \approx \mathbb{Z}/7\mathbb{Z}$$

Finally, this means that

$$\boxed{\mathbb{Z}[x]/(2x^2 - 4, 4x - 5) \approx \mathbb{Z}/7\mathbb{Z}}$$

(e)

Consider the homomorphism  $f : \mathbb{Z}[x] \rightarrow \mathbb{Z}[\sqrt{-3}]$  given by  $x \mapsto \sqrt{-3} = i\sqrt{3}$ . It is clearly surjective. We will find its kernel.

By definition  $p(x)$  is in  $\ker f$  if and only if  $f(p(x)) = 0$ . By definition of  $f$ , this is if and only if  $p(i\sqrt{3}) = 0$ . Thus,  $p(x)$  is in the kernel if and only if  $i\sqrt{3}$  is its root. By the complex conjugate theorem,  $-i\sqrt{3}$  is also a root of  $p(x)$ . Thus, the polynomial

$$(x - i\sqrt{3})(x + i\sqrt{3}) = x^2 + 3$$

divides  $p(x)$ , so

$$p(x) = (x^2 + 3)q(x)$$

By all of the above,  $p(x)$  is in the kernel of  $f$  if and only if it is of the above form; thus,

$$\ker f = (x^2 + 3)$$

By the First Isomorphism Theorem,

$$\mathbb{Z}/(x^2 + 3) \approx \mathbb{Z}[\sqrt{-3}]$$

Moreover, since the residue of 5 by  $f$  is 5,

$$\mathbb{Z}/(x^2 + 3, 5) \approx \mathbb{Z}[\sqrt{-3}]/(5)$$

Now define a homomorphism  $g : \mathbb{Z}[\sqrt{-3}] \rightarrow (\mathbb{Z}/5\mathbb{Z})[\sqrt{-3}]$  by

$$g(a + b\sqrt{-3}) = \bar{a} + \bar{b}\sqrt{-3}$$

It is clearly a surjective homomorphism, with the kernel  $(5)$  (because for  $a + b\sqrt{-3}$  to be in the kernel, both  $a$  and  $b$  must be divisible by 5). Thus, by the First Isomorphism Theorem,

$$\mathbb{Z}[\sqrt{-3}]/(5) \approx (\mathbb{Z}/5\mathbb{Z})[\sqrt{-3}]$$

Finally,

$$\boxed{\mathbb{Z}(x^2 + 3, 5) \approx (\mathbb{Z}/5\mathbb{Z})[\sqrt{-3}]}$$

## Result

(a)  $(\mathbb{Z}/2\mathbb{Z})[i]$

(b)  $\mathbb{Z}/5\mathbb{Z}$

(c)  $\mathbb{Z}/3\mathbb{Z}$

(d)  $\mathbb{Z}/7\mathbb{Z}$

(e)  $(\mathbb{Z}/5\mathbb{Z})[\sqrt{-3}]$

4. a



Suppose that they are isomorphic. We will first inspect some properties of these rings.

First of all, if we define a homomorphism with  $f(x) : \mathbb{Z}[x] \rightarrow \mathbb{Z}[\sqrt{-7}]$ ,  $x \mapsto \sqrt{-7}$ . It is clearly surjective, because  $f(a + bx) = a + b\sqrt{-7}$ , for all integers  $a, b$ .

Now let us find its kernel. Suppose that  $p(x) \in \ker f$ . Then  $f(p(x)) = 0$ , so  $p(\sqrt{-7}) = 0$ , which means that  $\sqrt{-7} = i\sqrt{7}$  is a root of  $p(x)$ . By the complex conjugate theorem,  $-i\sqrt{7}$  is also a root of  $p(x)$ . Thus, the polynomial

$$(x - i\sqrt{7})(x + i\sqrt{7}) = x^2 + 7$$

divides  $p(x)$ ; that is,

$$p(x) = (x^2 + 7)q(x)$$

Thus,  $p(x) \in (x^2 + 7)$ , so  $\ker f \subseteq (x^2 + 7)$ .

For the other inclusion, let  $h(x) \in (x^2 + 7)$ . Then  $h(x) = (x^2 + 7)g(x)$  for some  $\mathbb{Z}[x]$ . Clearly  $h(i\sqrt{7}) = 0$ , so  $f(h(x)) = 0$ , and  $h(x) \in \ker f$ .

Now we conclude that

$$\ker f = (x^2 + 7),$$

so, by the First Isomorphism Theorem,  $\mathbb{Z}[x]/(x^2 + 7)$  is isomorphic to  $\mathbb{Z}[\sqrt{-7}]$ .

Now we conclude that  $\mathbb{Z}[x]/(2x^2 + 7)$  is isomorphic to  $\mathbb{Z}[\sqrt{-7}]$ . We will prove that this is impossible.

First of all,

$$(2 + (2x^2 + 7))((x^2 + 4) + (2x^2 + 7)) = 1 + (2x^2 + 7),$$

so  $2 + (2x^2 + 7)$  is invertible in  $\mathbb{Z}[x]/(2x^2 + 7)$ .

On the other hand, we will prove that  $2$  is not invertible in  $\mathbb{Z}[\sqrt{-7}]$ . Suppose that it is. Then there exists some  $a + b\sqrt{-7} \in \mathbb{Z}[\sqrt{-7}]$ , with  $a, b$  integers, such that

$$2(a + b\sqrt{-7}) = 1$$

This means that

$$2a + 2b\sqrt{-7}i = 1$$

and

$$2a = 1,$$

which is absurd. Thus,  $2$  is not invertible in  $\mathbb{Z}[\sqrt{-7}]$ .

Now let  $\varphi : \mathbb{Z}[x]/(2x^2 + 7) \rightarrow \mathbb{Z}[\sqrt{-7}]$  be some isomorphism. First of all, by definition of homomorphisms,  $\varphi(1 + (2x^2 + 7)) = 1$  and

$$\varphi(2 + (2x^2 + 7)) = \varphi(1 + (2x^2 + 7)) + \varphi(1 + (2x^2 + 7)) = 1 + 1 = 2$$

Now let  $u = (2 + (2x^2 + 7))^{-1}$ . Then

$$1 = \varphi(1 + (2x^2 + 7)) = \varphi((2 + (2x^2 + 7))u) = \varphi(2 + (2x^2 + 7))\varphi(u) = 2\varphi(u)$$

However, this means that  $\varphi(u) = 2^{-1}$  in  $\mathbb{Z}[\sqrt{-7}]$ , which is impossible (as proven before).

Thus,  $\mathbb{Z}[x]/(2x^2 + 7)$  and  $\mathbb{Z}[\sqrt{-7}]$  are not isomorphic. With this, we conclude that  $\mathbb{Z}[x]/(2x^2 + 7)$  and  $\mathbb{Z}[x]/(x^2 + 7)$  are not isomorphic.

## Result

3 of 3

They are not isomorphic.

## Section 5

1. a

We use the result of Proposition 11.5.5 (c).

Let  $\beta_1 = \alpha^3 + \alpha^2 + \alpha$ ,  $\beta_2 = \alpha^5 + 1$ . Then  $g_1(x) = x^3 + x^2 + x$ ,  $g_2(x) = x^5 + 1$ , and

$$g_1(x)g_2(x) = x^8 + x^7 + x^6 + x^3 + x^2 + x$$

Now we need to divide  $g_1(x)g_2(x)$  by  $f(x)$ .

$$\begin{array}{r} x^4 \phantom{+ x^3 + x^2 + x + 1} \phantom{+ x^3 + x^2 + x} \\ x^4 + x^3 + x^2 + x + 1 \overline{) x^8 + x^7 + x^6 + x^3 + x^2 + x} \\ \underline{-x^8 - x^7 - x^6 - x^5 - x^4} \phantom{+ x^3 + x^2 + x} \\ -x^5 - x^4 + x^3 + x^2 + x \\ \underline{x^5 + x^4 + x^3 + x^2 + x} \\ 2x^3 + 2x^2 + 2x \end{array}$$

Thus,

$$g_1(x)g_2(x) = f(x)(x^4 - x) + 2x^3 + 2x^2 + 2x$$

Thus,

$$\boxed{\beta_1\beta_2 = 2\alpha^3 + 2\alpha^2 + 2\alpha}$$

## Result

$$2\alpha^3 + 2\alpha^2 + 2\alpha$$

2. a

The relation is

$$\alpha - a = 0$$

Thus, what we really need to show is that

$$R[x]/(x - a) \approx R$$

Observe the homomorphism  $\mathbb{R}[x] \rightarrow R$  which sends  $x \rightsquigarrow a$ . Then its kernel is  $(x - a)$ , and it is clearly isomorphic (for any  $r \in R$ , we have that  $x + (r - a) \rightsquigarrow r$ ). Thus, by the First Isomorphism Theorem,

$$R[x]/(x - a) \approx R$$

as required.

## Result

What we really need to show is

$$R[x]/(x - a) \approx R$$

(Hint: use the First Isomorphism Theorem.)

## 3. a

The inverse of 2 is a number  $\alpha$  such that

$$2\alpha = 1$$

Thus, we get the relation

$$2\alpha - 1 = 0$$

Therefore, we need to identify the ring

$$(\mathbb{Z}/12\mathbb{Z})[x]/(2x - 1)$$

Let  $I = (2x - 1)$ . Since  $12 = 0$  in  $\mathbb{Z}/12\mathbb{Z}$ , we get that

$$6 = 6 - 12x = (-6)(2x - 1)$$

Therefore,  $6 \in I$ . Now,

$$6x - 3 = 3(2x - 1)$$

is in  $I$ . Furthermore,

$$3 = 6x - 3(2x - 1)$$

Thus,  $3 \in I$ .

Furthermore,

$$x - 2 = -(2x - 1) + 3(x - 1)$$

Thus,  $x - 2 \in I$ .

Finally,  $x - 2$  is monic in  $(\mathbb{Z}/12\mathbb{Z})[x]$ , so we can divide  $f(x) \in (\mathbb{Z}/12\mathbb{Z})[x]$  by it:

$$f(x) = q(x)(x - 2) + m,$$

where  $m \in \mathbb{Z}/12\mathbb{Z}$ . Furthermore, we can write  $m = 3k + l$ , for some  $l \in \{0, 1, 2\}$ . We conclude that

$$f(x) = q(x)(x - 2) + 3k + l \implies f(x) - l = q(x)(x - 2) + 3k$$

The right side is in  $I$ , so  $f(x) - l \in I$ . This means that every element of the quotient ring  $(\mathbb{Z}/12\mathbb{Z})[x]/I$  is equal to  $l + I$ , where  $l \in \{0, 1, 2\}$ . This precisely means that

$$(\mathbb{Z}/12\mathbb{Z})[x]/(2x - 1) \approx \mathbb{Z}/3\mathbb{Z}$$

## Result

2 of 2

We get a ring isomorphic to  $\mathbb{Z}/3\mathbb{Z}$ .

4. a

(a)

We need to identify the ring

$$\mathbb{Z}[x]/(2x - 6, 6x - 15)$$

Let  $I = (2x - 6, 6x - 15)$ . Notice that

$$3 = (2x - 6)(-3) + (6x - 15),$$

so  $3 \in I$ . Moreover,

$$x = 3(x - 2) + (2x - 6) \cdot (-1)$$

Thus,  $x \in I$ .

Now we will show that

$$I = (3, x)$$

Since  $3, x \in I$ , and  $(3, x)$  is the smallest ideal containing these elements, we conclude that

$$(3, x) \subseteq I$$

On the other hand,  $2x - 6$  and  $6x - 15$  are clearly in  $(3, x)$ , so we have that

$$I = (2x - 6, 6x - 15) \subseteq (3, x)$$

Finally,

$$I = (3, x)$$

Thus, we identify the ring  $\mathbb{Z}[x]/(3, x)$ .

First of all, consider the homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z}$  which sends  $x \mapsto 0$ . Clearly it is surjective and its kernel is  $(x)$ .

Thus, by the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(x) \approx \mathbb{Z}$$

The residue of 3 is 3, so we kill 3 in  $\mathbb{Z}$  and get

$$\mathbb{Z}[x]/(x, 3) = \mathbb{Z}/3\mathbb{Z}$$

(b)

We need to identify the ring

$$\mathbb{Z}[x]/(2x - 6, x - 10)$$

Let  $I = (2x - 6, x - 10)$ . Notice that

$$14 = (2x - 6) + (x - 10)(-2),$$

so  $14 \in I$ .

Now we will show that

$$I = (14, x - 10)$$

Since  $14, (x - 10) \in I$ , and  $(14, x - 10)$  is the smallest ideal containing these elements, we conclude that

$$(14, x - 10) \subseteq I$$

On the other hand,

$$2x - 6 = (x - 10) \cdot 2 + 14$$

which means that  $2x - 6 \in I$ . Thus,

$$I = (2x - 6, x - 10) \subseteq (14, x - 10)$$

Finally,

$$I = (14, x - 10)$$

Thus, we identify the ring  $\mathbb{Z}[x]/(14, x - 10)$ .

First of all, consider the homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z}$  which sends  $x \rightsquigarrow 10$ . Clearly it is surjective and its kernel is  $(x - 10)$ . Thus, by the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(x - 10) \approx \mathbb{Z}$$

The residue of 14 is 14, so we kill 14 in  $\mathbb{Z}$  and get

$$\boxed{\mathbb{Z}[x]/(14, x - 10) = \mathbb{Z}/14\mathbb{Z}}$$

(NOTE: since we already had a monic polynomial, we could immediately use the homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z}$  which sends  $x \rightsquigarrow 10$  and confirm that the residue of  $2x - 6$  is 14.)

(c)

We need to identify the ring

$$\mathbb{Z}[x]/(x^3 + x^2 + 1, x^2 + x)$$

Let  $I = (x^3 + x^2 + 1, x^2 + x)$ . Then

$$1 = (x^3 + x^2 + 1) + (x^2 + x)(-x)$$

So,  $1 \in I$ . But this means that  $I = \mathbb{Z}[x]$ ! Moreover, this means that

$$\boxed{\mathbb{Z}[x]/I = (0)}$$

(the zero ring).

### Result

(a)  $\mathbb{Z}/3\mathbb{Z}$

(b)  $\mathbb{Z}/14\mathbb{Z}$

(c)  $(0)$  (the zero ring)

5. a

Let  $I = (x^2)$ ,  $J = (x^2 - 1)$ . If  $F$  is of characteristic 2, then

$$(x - 1)^2 = x^2 - 2x + 1 = x^2 + 1 = x^2 - 1$$

(the last equality follows from  $1 = -1$  because of characteristic 2). Now define a homomorphism  $F[x] \rightarrow F[x]/((x - 1)^2)$  which sends  $x \mapsto x - 1$ . Then this ring is clearly surjective and its kernel is  $(x^2)$ . Thus, by the First Isomorphism Theorem,

$$F[x]/(x^2) \approx F[x]/((x - 1)^2)$$

Now suppose that  $F$  is not of characteristic 2 (that is, its characteristic is more than 2). Here we have that  $-1 \neq 1$  (because this would imply that  $2 = 0$  in  $F$ , and  $2r = 0$  for every  $r \in F$ , so  $F$  would be of characteristic 2).

Now we see that

$$(x + (I))(x + (I)) = x^2 + I = I,$$

since  $x^2 \in I$ .

On the other hand, we now suppose that there exists some polynomial  $f(x) \in F[x]$  such that

$$(f(x) + J)(f(x) + J) = J$$

This means that

$$f(x)^2 + J = J,$$

so  $f(x)^2 \in J$ . Since  $J = (x^2 - 1)$ , this means that

$$f(x)^2 = (x^2 - 1)g(x)$$



for some polynomial  $g(x) \in F[x]$ . From this we get that

$$f(1)^2 = (1 - 1)g(1) = 0$$

and

$$f(-1)^2 = 0$$

Now let  $f(1) = a$ , for some  $a \in F$ . Then  $a^2 = 0$ , and, after multiplying this equality by  $a^{-1}$ , we get  $a = 0$ . Thus,  $f(1) = 0$ . Similarly we get that  $f(-1) = 0$ .

So, 1 and  $-1$  are roots of  $f(x)$ , meaning that  $x - 1$  and  $x + 1$  divide  $f(x)$ . Thus,

$$f(x) = (x + 1)(x - 1)p(x) = (x^2 - 1)p(x)$$

for some polynomial  $p(x) \in F[x]$ . But this means that  $f(x) \in J$ , so

$$f(x) + J = J$$

This means that  $F[x]/J$  does not have any nonzero nilpotent elements.

Suppose that  $F[x]/I$  and  $F[x]/J$  are isomorphic, and let  $\varphi$  be some isomorphism. Let  $p(x) + J = \varphi(x + I)$ . Then

$$p(x)^2 + J = \varphi(x^2 + I) = \varphi(I) = J$$

Thus,  $p(x)^2 + J = J$ , which means that  $p(x) + J = J$ . But this means that  $x + I$  is in the kernel of  $\varphi$  (since  $J$  is zero in  $F[x]/J$ ). Thus, the kernel of  $\varphi$  contains something other than a zero (here the zero is  $I$ ), which is impossible since  $\varphi$  is an isomorphism so it must be injective and have a trivial kernel!

Thus, if  $F$  has a characteristic of more than 2, the rings are not isomorphic.

## Result

3 of 3

They are isomorphic if and only if  $F$  is of characteristic 2.

6. a

(a)

Since  $\beta \in R'$ , it is of the form

$$\beta = b_0 + b_1\alpha + \dots + b_k\alpha^k$$

where  $b_0, b_1, \dots, b_k \in R$ .

Since  $a\alpha = \alpha a = 1$ , we can write

$$\beta = \alpha^k(a^k b_0 + a^{k-1} b_1 + \dots + b_k)$$

Since  $a^k b_0 + a^{k-1} b_1 + \dots + b_k \in R$ , we have proven the statement.

(b)

Denote the map by  $\varphi$ . Suppose that  $a^n b = 0$ . Then

$$\varphi(a^n b) = \varphi(0) = 0$$

Moreover,

$$\varphi(a^n b) = \varphi(a)^n \varphi(b) = a^n \varphi(b)$$

Thus,

$$a^n \varphi(b) = 0$$

Since  $a$  has an inverse  $\alpha$  (in the ring  $R'$  of course), we multiply this equality by  $\alpha^n$  to get

$$\varphi(b) = 0$$

Thus,  $b$  is in the kernel of  $\varphi$ .

On the other hand, we use the fact that  $R' = R[x]/(ax - 1)$  to conclude that if  $b$  is in the kernel of  $\varphi$ , then  $b \in (ax - 1)$  (because the kernel of the whole projective mapping  $\pi$  is  $(ax - 1)$ ). From this,

$$b = (ax - 1)p(x)$$

for some polynomial  $p(x) \in R[x]$ .

Now we write

$$p(x) = c_n x^n + \dots + c_1 x + c_0$$

Thus,

$$b = (ax - 1)p(x) = (ac_n)x^{n+1} + (ac_{n-1} - c_n)x^n + \dots + (ac_0 - c_1)x + (-c_0)$$

This is a polynomial equation, so

$$\begin{aligned} b &= -c_0 \\ ac_0 - c_1 &= 0 \\ &\vdots \\ ac_{n-1} - c_n &= 0 \\ ac_n &= 0 \end{aligned}$$

Now we get equalities:

$$\begin{aligned} b &= -c_0 \\ ab &= -ac_0 = -c_1 \\ a^2 b &= -ac_1 = -c_2 \\ &\vdots \\ a^n b &= -ac_{n-1} = -c_n \\ a^{n+1} b &= -ac_n = 0 \end{aligned}$$

Thus,  $a^{n+1}b = 0$ , so the statement is proven.

(c)

Suppose that  $a$  is nilpotent; thus is, there exists some positive integer  $n$  such that  $a^n = 0$ . Now we have that  $a^n b = 0$ , for all  $b \in R$ , so all elements of  $R$  are in the kernel of  $\varphi$ ! This means that  $R'$  must be a zero ring. (Otherwise, by the definition of a homomorphism, we would have  $\varphi(1) = 1$ , which is impossible here.)

On the other hand, suppose that  $R'$  is a zero ring. Then we must have that  $\varphi(b) = 0$  for all  $b \in R$  (since  $\varphi(b) \in R' = \{0\}$ ). Specially, 1 is in the kernel of  $\varphi$ . This means that there exists some positive integer  $n$  such that  $a^n \cdot 1 = 0$ . However,  $a^n \cdot 1 = a^n$ , so  $a^n = 0$ , meaning that  $a$  is nilpotent.

## Result

4 of 4

(a) Use the fact that  $\beta = b_0 + b_1\alpha + \dots + b_n\alpha^n$  for some  $b_i \in R$  and that  $a\alpha = \alpha a = 1$ .

(b) It is easier when you write  $R' = R[x]/(ax - 1)$ .

(c) Use (b).

7. a

By definition, Laurent polynomials are polynomials of the form

$$f(t) = \sum_{i=-m}^n a_i t^i = a_{-m} t^{-m} + \dots + a_0 + \dots + a_n t^n$$

for some nonnegative integers  $m, n$  and  $a_i \in F$ .

Now we denote by  $L$  the set of Laurent polynomials, and observe a homomorphism  $R[x] \rightarrow L$  which sends  $x \rightsquigarrow t^{-1}$ . Since  $g(x) \in R[x]$  (because  $R = F[t]$ ) is of the form

$$g(x) = a_0(t) + a_1(t)x + \dots + a_n(t)x^n$$

(so coefficients are polynomials in variable  $t$ ), we will prove that this homomorphism is surjective. Let  $f(t) \in L$ ,

$$f(t) = \sum_{i=-m}^n a_i t^i$$

and

$$g(x) = (a_{-m} + a_{-m+1}t + \dots + a_n t^{m+n})x^m$$

Now we see that  $g(x) \rightsquigarrow f(t)$ , so the homomorphism is really surjective.

Now we find its kernel. Suppose that  $g(x)$  is in the kernel. Divide  $g(x)$  by  $(tx - 1)$ :

$$g(x) = q(x)(tx - 1) + r(x)$$

Thus,  $r(x)$  is a constant polynomial. Moreover,  $g(t^{-1})$  yields

$$r(t^{-1}) = 0.$$

so  $r(x)$  is a zero polynomial. Finally,

$$g(x) = q(x)(tx - 1)$$

This means that the kernel is the set  $(tx - 1)$ .

Finally, be the First Isomorphism Theorem,

$$R[x]/(tx - 1) \approx L$$

as required.

## Result

Prove that

$$R[x]/(tx - 1) \approx L$$

by defining a homomorphism  $R[x] \rightarrow L$  which sends  $x \mapsto t^{-1}$ .

## Section 6

1. a

First of all, by the definition of  $\varphi$ , we get that

$$\varphi(f(x)) = (f(1), f(i))$$

Now let  $f(x)$  be in the kernel of  $\varphi$ . Then

$$(0, 0) = \varphi(f(x)) = (f(1), f(i))$$

This means that  $f(1) = f(i) = 0$ , so 1 and  $i$  are roots of  $f(x)$ . Furthermore, by the complex conjugate theorem we conclude that  $-i$  is also a root of  $f(x)$ . This means that polynomials  $x - 1$ ,  $x - i$ , and  $x + i$  divide  $f(x)$ .

Since

$$(x - 1)(x - i)(x + i) = (x - 1)(x^2 + 1),$$

so  $(x - 1)(x^2 + 1)$  divides  $f(x)$ . This means that

$$f(x) = (x - 1)(x^2 + 1)q(x)$$

for some polynomial  $q(x) \in \mathbb{R}[x]$ . This means that  $f(x) \in ((x - 1)(x^2 + 1))$ , so

$$\ker \varphi \subseteq ((x - 1)(x^2 + 1))$$

For the other inclusion, let  $g(x) \in ((x - 1)(x^2 + 1))$ . Then  $g(x) = h(x)(x - 1)(x^2 + 1)$  for some  $h(x) \in \mathbb{R}[x]$ . Now  $g(1) = g(i) = 0$ , so  $\varphi(g(x)) = (0, 0)$ , and  $g(x) \in \ker \varphi$ . Therefore,

$$((x - 1)(x^2 + 1)) \subseteq \ker \varphi$$

and

$$\ker \varphi = ((x - 1)(x^2 + 1))$$

Now to find the image. First of all,  $f(1) \in \mathbb{R}$ , for every polynomial  $f(x) \in \mathbb{R}[x]$ , so

$$\text{im}\varphi \subseteq \mathbb{R} \times \mathbb{C}$$

We will prove that the other inclusion also holds. Let

$$(a, b + ci) \in \mathbb{R} \times \mathbb{C}$$

Now we need to find a polynomial  $f(x) \in \mathbb{R}[x]$  such that

$$\varphi(f(x)) = (a, b + ci)$$

That is,

$$(f(1), f(i)) = (a, b + ci)$$

Since coefficients of  $f(x)$  are real, this will become a system of three equations. Thus, we need three variables to solve it. Let  $f(x) = d_0 + d_1x + d_2x^2$ . Then

$$\varphi(f(x)) = (f(1), f(i)) = (d_0 + d_1 + d_2, d_0 + d_1i - d_2)$$

Thus,

$$\begin{aligned} d_0 + d_1 + d_2 &= a \\ d_0 - d_2 &= b \\ d_1 &= c \end{aligned}$$

Thus,  $d_1 = c$  Now the system becomes (ignoring the third equation because we do not need it anymore)

$$\begin{aligned} d_0 + d_2 &= a - c \\ d_0 - d_2 &= b \end{aligned}$$

Adding these equations together, we get

$$d_0 = \frac{a + b - c}{2} \quad \text{Similarly, by subtracting the second equation from the first, we get } d_2 = \frac{a - b - c}{2}$$

Thus, when we set  $f(x) = d_0 + d_1x + d_2x^2$  with these  $d_i$ , we get

$$\varphi(f(x)) = (a, b + ci)$$

So, every point of  $\mathbb{R} \times \mathbb{C}$  gets hit, so

$$\mathbb{R} \times \mathbb{C} \subseteq \text{im}\varphi$$

and

$$\text{im}\varphi = \mathbb{R} \times \mathbb{C}$$

## Result

$$\ker\varphi = ((x-1)(x^2+1)), \quad \text{im}\varphi = \mathbb{R} \times \mathbb{C}$$

2. a

### $\mathbb{Z}/(6)$ and $\mathbb{Z}/(2) \times \mathbb{Z}/(3)$

We will prove that the statement holds. First,  $\mathbb{Z}/(2) \times \mathbb{Z}/(3)$  is equal to the set

$$\{((2), (3)), ((2), 1 + (3)), ((2), 2 + (3)), (1 + (2), (3)), (1 + (2), 1 + (3)), (1 + (2), 2 + (3))\}$$

Define a mapping

$$\varphi : \mathbb{Z} \rightarrow \mathbb{Z}/(2) \times \mathbb{Z}/(3), \quad \varphi(m) = (m + (2), m + (3))$$

Now we prove that  $\varphi$  is surjective. Truly,

$$\varphi(0) = ((2), (3))$$

$$\varphi(1) = (1 + (2), 1 + (3))$$

$$\varphi(2) = (2 + (2), 2 + (3)) = ((2), 2 + (3))$$

$$\varphi(3) = (3 + (2), 3 + (3)) = (1 + (2), (3))$$

$$\varphi(4) = (4 + (2), 4 + (3)) = ((2), 1 + (3))$$

$$\varphi(5) = (5 + (2), 5 + (3)) = (1 + (2), 2 + (3))$$

Now we need to find a kernel of  $\varphi$ . Suppose that  $m$  is in the kernel of  $\varphi$ . Then, because  $(2)$  is the zero in  $\mathbb{Z}/(2)$  and  $(3)$  is the zero in  $\mathbb{Z}/(3)$ ,

$$((2), (3)) = \varphi(m) = (m + (2), m + (3))$$

Thus,  $m + (2) = (2)$ , which means that  $m \in (2)$ , so there exists some integer  $k$  such that  $m = 2k$ . Furthermore,  $m + (3) = (3)$ , so  $m \in (3)$ . This means that  $2k \in (3)$ , so  $2k = 3l$ , for some integer  $l$ . This means that 3 divides  $2k$ . Moreover, 3 is prime, so it must divide either 2 or  $k$ . Since it does not divide 2, it divides  $k$ ; thus,  $k = 3n$ , for some integer  $n$ . Finally,

$$m = 2k = 6n$$

for some integer  $n$ , so

$$m \in (6)$$

This means that

$$\ker \varphi \subseteq (6)$$

To prove the other inclusion, let  $a \in (6)$ . Then  $a = 6b$ , for some integer  $b$ . Thus,

$$\varphi(a) = \varphi(6b) = (6b + (2), 6b + (3)) = ((2), (3))$$

since  $6b \in (2)$  and  $6b \in (3)$ , so  $6b + (2) = (2)$  and  $6b + (3) = (3)$ . So,  $a \in \ker \varphi$ . Finally, this means that

$$(6) \subseteq \ker \varphi$$

and

$$\ker \varphi = (6)$$

By the First Isomorphism Theorem, we conclude that

$$\mathbb{Z}/(6) \approx \mathbb{Z}/(2) \times \mathbb{Z}/(3)$$



Let  $\varphi : \mathbb{Z}/(8) \rightarrow \mathbb{Z}/(2) \times \mathbb{Z}/(8)$  be some homomorphism. Then we must have that  $\varphi(1 + (8)) = (1 + (2), 1 + (4))$  (because homomorphism sends the multiplicative identity to the multiplicative identity). Now, because  $\varphi$  is a homomorphism

$$\begin{aligned}\varphi(4 + (8)) &= \varphi((1 + (8)) + (1 + (8)) + (1 + (8)) + (1 + (8)) + (1 + (8))) \\ &= \varphi(1 + (8)) + \varphi(1 + (8)) + \varphi(1 + (8)) + \varphi(1 + (8)) \\ &= (1 + (2), 1 + (4)) + (1 + (2), 1 + (4)) + (1 + (2), 1 + (4)) + (1 + (2), 1 + (4)) \\ &= (4 + (2), 4 + (4)) \\ &= ((2), (4))\end{aligned}$$

(the last equality holds because  $4 \in (2)$  and  $4 \in (4)$ , so  $4 + (2) = (2)$  and  $4 + (4) = (4)$ ).

Thus,  $4 + (8) \in \ker \varphi$ . However,  $4 + (8)$  is not a zero in  $\mathbb{Z}/(8)$ ! Thus,  $\varphi$  has a kernel which contains a nonzero element, so it cannot be injective. Neither can it be isomorphism. Since  $\varphi$  was arbitrarily taken homomorphism, we conclude that there exists no isomorphism between these two rings, so they are not isomorphic.

**SOL**

First: yes

Second: no

3. a

Let  $R$  be the ring of order 10. Let  $R^+$  be the additive group  $(R, +)$ .

Since  $10 = 2 \cdot 5$ , by the First Sylow Theorem, there exists a  $p$ -subgroup of order 5 in  $R^+$ . By the Third Sylow Theorem, the number of such subgroup is  $5k + 1$ , where  $k$  is a nonnegative integer and it divides the order of the group  $R^+$ : 10. This means that we must have only one  $p$ -group; by the Corollary of the Second Sylow Theorem, we conclude that it is a normal subgroup of  $R^+$ , so we denote it by  $N$ .

Since it is normal, it makes sense to observe a quotient group  $R^+/N$ ; since

$$|R^+/N| = \frac{|R^+|}{|N|} = 2$$

Thus,

$$R^+/N = \{N, aN\}$$

for some  $a \in R^+, a \notin N$ . This means that

$$R^+ = N \cup aN$$

so

$$R^+ = \{0, b, 2b, 3b, 4b, ab, 2ab, 3ab, 4ab\}$$

(here we write the elements using the additive notation; thus  $b^k = kb$ ).

Moreover, now it is clear that

$$R^+ = H + N,$$

where  $H = \{0, a\}$ . Furthermore, since  $R^+$  is commutative, both  $H$  and  $N$  are normal in  $R^+$ . Finally,  $H \cap N = \{0\}$ . This means that

$$R^+ \approx H \times N$$

Now we easily see that  $H \approx \mathbb{Z}/2\mathbb{Z}$  and  $N \approx \mathbb{Z}/5\mathbb{Z}$ . Furthermore, a function

$$f : (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/5\mathbb{Z}) \rightarrow \mathbb{Z}/10\mathbb{Z}$$

given by

$$f(a, b) = (5a) + b$$

is easily seen to be an isomorphism (of additive groups).

Thus,

$$R^+ \approx H \times N \approx (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/5\mathbb{Z}) \approx \mathbb{Z}/10\mathbb{Z}$$

Since  $R^+$  is isomorphic to  $\mathbb{Z}/10\mathbb{Z}$ , we will prove that our ring  $R$  is also isomorphic to  $\mathbb{Z}/10\mathbb{Z}$ .

Let  $\bar{1}$  be a multiplicative identity in  $R$ . We will prove that it is of order 10 in  $R^+$ . If it is of order less than 10, say  $n$ , then

$$nr = (n \cdot \bar{1})r = 0$$

Thus, all elements of  $R^+$  are of order less than 10, which is impossible since  $R^+$  is isomorphic to  $\mathbb{Z}/10\mathbb{Z}$ , so it must have an element of order 10 (since  $\mathbb{Z}/10\mathbb{Z}$  has an element of order 10).

So,  $\bar{1}$  is of order 10, and since  $R$  also has 10 elements, we conclude that

$$R = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \dots, \bar{9}\},$$

where  $\bar{2} = 1 + 1$  and so on. Now it is easily seen that

$$\varphi : R \rightarrow \mathbb{Z}/10\mathbb{Z}, \quad \varphi(\bar{a}) = a$$

is an isomorphism (of rings!).

The fact that  $\varphi(\bar{1}) = 1$  is clear from the definition. Moreover,  $\varphi(\bar{a} + \bar{b}) = \varphi(\bar{a}) + \varphi(\bar{b})$  follows because  $\bar{1}$  is of order 10 (so, the same as 1 in  $\mathbb{Z}/10\mathbb{Z}$ ), and

$$\bar{a} = \underbrace{\bar{1} + \dots + \bar{1}}_{a \text{ times}}$$

Similarly, because of the distributive law, the product of  $\bar{a}$  and  $\bar{b}$  is the sum of  $ab$   $\bar{1}$ 's, which proves that

$$\varphi(\bar{a}\bar{b}) = \varphi(\bar{a})\varphi(\bar{b})$$

## Result

4 of 4

All of them are isomorphic to  $\mathbb{Z}/10\mathbb{Z}$ .

4. a

**(a)**

This relation is given by a monic polynomial  $f(x) = x^2 + x + 1$ , so  $\{1, \alpha\}$  is a basis for  $\mathbb{F}_2[\alpha]$  (see Proposition 11.5.5 (a)). Thus,

$$R[\alpha] = \{a + b\alpha \mid a, b \in \mathbb{F}_2\} = \{0, 1, \alpha, 1 + \alpha\}$$

Furthermore,

$$\alpha(1 + \alpha) = \alpha + \alpha^2 = -1 = 1$$

(from relation, and  $-1 = 1$  in  $\mathbb{F}_2$ ), so  $\alpha$  and  $1 + \alpha$  are invertible in  $R[\alpha]$ . This means that  $R[\alpha]$  is a field of four elements! (1 is clearly invertible.)

**(b)**

The same as in **(a)**,

$$R[\alpha] = \{0, 1, \alpha, 1 + \alpha\}$$

Since

$$\alpha^2 = -1 = 1,$$

$\alpha$  is invertible (recall that  $-1 = 1$  in  $\mathbb{F}_2$ ). Furthermore,

$$(1 + \alpha)^2 = 1 + 2\alpha + \alpha^2 = 1 + \alpha^2 = 0$$

Thus, it cannot be invertible. If it was invertible, multiplying the above equation by  $(1 + \alpha)^{-1}$  would yield  $1 + \alpha = 0$  which is impossible.

Thus, in this case, this set is not a field. Also, we cannot say anything more which would be interesting.

**(c)**

As before,

$$R[\alpha] = \{0, 1, \alpha, 1 + \alpha\}$$

Here

$$\alpha^2 = -\alpha = \alpha$$

(since  $-\alpha = \alpha$  in  $\mathbb{F}_2$ ), so  $\alpha$  is idempotent. By Proposition 11.6.2, we get that

$$R[\alpha] \approx (\alpha R) \times ((1 - \alpha)R)$$

First of all,

$$\alpha R = \{0, \alpha, \alpha^2, \alpha^2 + \alpha\} = \{0, \alpha\}$$

after using that  $\alpha^2 = \alpha$ . Moreover,  $1 - \alpha = 1 + \alpha$ , and

$$(1 + \alpha)R = \{0, 1 + \alpha, \alpha + \alpha^2, (1 + \alpha)^2\} = \{0, \alpha\}$$

since  $(1 + \alpha)^2 = 1 + 2\alpha + \alpha^2 = 1 + \alpha$ .

Thus, since  $\alpha^2 = \alpha$ , so it will serve as an identity,

$$\alpha R \approx \mathbb{F}_2$$

$$(1 - \alpha)R \approx \mathbb{F}_2$$

Finally,

$$\boxed{R[\alpha] \approx \mathbb{F}_2 \times \mathbb{F}_2}$$

## Result

All sets are  $R[\alpha] = \{0, 1, \alpha, 1 + \alpha\}$ .

(a) Here  $R[\alpha]$  is a field.

(b) Here  $R[\alpha]$  is not a field, but  $\alpha$  is invertible ( $1 + \alpha$  is not).

(c) Here  $R[\alpha]$  is isomorphic to  $\mathbb{F}_2 \times \mathbb{F}_2$ .

## 5. a

The relation is

$$\alpha^2 - 1 = 0$$

Thus, we must identify the ring

$$\mathbb{R}[x]/(x^2 - 1)$$

Define  $I = (x^2 - 1) = ((x - 1)(x + 1))$  (since  $x^2 - 1 = (x - 1)(x + 1)$ ). Now we will define a function

$$f : \mathbb{R}[x] \rightarrow \mathbb{R} \times \mathbb{R}, \quad f(p(x)) = (p(1), p(-1))$$

This is clearly a homomorphism. Now we will find its kernel and image.

Suppose that  $p(x) \in \mathbb{R}[x]$  is in the kernel of  $f$ . Then

$$(0, 0) = f(p(x)) = (p(1), p(-1))$$

Thus,

$$p(1) = p(-1) = 0,$$

so  $p(x)$  has roots 1 and  $-1$ . This means that polynomials  $(x - 1)$  and  $(x + 1)$  divide  $p(x)$ . Moreover,

$$(x - 1)(x + 1) = x^2 - 1,$$

so  $(x^2 - 1)$  divides  $p(x)$ ; that is,

$$p(x) = (x^2 - 1)q(x)$$

for some polynomial  $q(x) \in \mathbb{R}[x]$ . Thus,  $p(x) \in (x^2 - 1) = I$ , and

$$\ker f \subseteq I$$

To prove the other inclusion, let  $p(x) \in I$ . Then  $p(x) = (x^2 - 1)q(x)$ , so  $p(1) = p(-1) = 0$ . Now

$$f(p(x)) = (p(1), p(-1)) = (0, 0),$$

so  $p(x) \in \ker f$ . Thus

$$I \subseteq \ker f$$

and

$$\ker f = I$$

Now we want to prove that the image of  $f$  is  $\mathbb{R} \times \mathbb{R}$ . Let  $(a, b) \in \mathbb{R} \times \mathbb{R}$ . We want to find a polynomial  $p(x) \in \mathbb{R}[x]$  such that

$$f(p(x)) = (a, b)$$

Since  $f(p(x)) = (p(1), p(-1))$ , this yields a system

$$\begin{aligned} p(1) &= a \\ p(-1) &= b \end{aligned}$$

So, we have two equations, so it would be great if we had two variables. Thus,  $p(x)$  would be of the first degree:

$$p(x) = \alpha + \beta x$$

Now our system becomes

$$\begin{aligned} \alpha + \beta &= a \\ \alpha - \beta &= b \end{aligned}$$

Adding these equation together, we get  $\alpha = \frac{a+b}{2}$ . Plugging this into the first equation yields  $\beta = \frac{a-b}{2}$ .

Thus,

$$f(p(x)) = (a, b)$$

when we set these coefficients, so  $(a, b)$  is in the image of  $f$ . Thus,  $f$  is surjective.

Now by the First Isomorphism Theorem we conclude that

$$\mathbb{R}[x]/(x^2 - 1) \approx \mathbb{R} \times \mathbb{R},$$

as required.

## Result

Hint: define the homomorphism

$$f : \mathbb{R}[x] \rightarrow \mathbb{R} \times \mathbb{R}, \quad f(p(x)) = (p(1), p(-1))$$

and use the First Isomorphism Theorem.

6. a

A set  $R$  is defined as the ring which shows the binary functions with respect to addition and multiplication and also satisfies the statement that the set is abelian under addition, monoid under multiplication and is distributive under multiplication with respect to addition.

[Comment](#)

Step 2 of 5 ^

To find: The ring obtained from the product ring  $\mathbb{R} \times \mathbb{R}$  by inverting the elements  $(2, 0)$

For the above proof first prove that the any adjoin element  $\alpha$  which satisfies the relation  $\alpha^2 = 1$  to the real number  $\mathbb{R}$  then the resulting ring is isomorphic to the product ring  $\mathbb{R} \times \mathbb{R}$

The proof is as follows;

For the proving the result first consider the resulting ring to be  $R$  such that;

$$R = \mathbb{R}[x]/(x^2 - 1)$$

Now, consider the map;

$$\partial: \mathbb{R}[x] \rightarrow \mathbb{R} \times \mathbb{R}$$

Defined as,

$$f \mapsto (f(1), f(-1))$$

So, from the above mapping the kernel will be;

$$\begin{aligned} \ker \partial &= (x^2 - 1) \\ &= ((x+1)(x-1)) \end{aligned}$$

Now, by first isomorphism theorem which states that;

Let  $f: R \rightarrow R'$  be a ring homomorphism with kernel  $K$  and let  $I$  be another ideal. Let  $\pi: R \rightarrow \bar{R}$  be the canonical map from  $R$  to  $\bar{R} = R/I$  and consider  $\bar{f}: \bar{R} \rightarrow R'$  such that  $\bar{f}\pi = f$ . If  $f$  is surjective and  $I = K$  then  $\bar{f}$  is an isomorphism

Hence, by the above theorem  $\alpha$  is surjective.

Thus,  $R \rightarrow \mathbb{R} \times \mathbb{R}$  is an isomorphism

[Comment](#)

Step 4 of 5 ^

Now, let;

$$\begin{aligned} f &= ax + b \\ &\in \mathbb{R}[x] \end{aligned}$$

Consider  $f$  to be of degree 1, then;

$$\partial(ax + b) = (a + b, a - b)$$

And, finally consider;

$$(a + b, a - b) = (x, y)$$



Hence, from the above proved isomorphism;

$$x+1 \rightarrow (2,0)$$

Such that the ring is given as shown below:

$$\begin{aligned} \frac{\mathbb{R}[x,y]}{(x^2-1, (x+1)y-1)} &\approx \frac{\mathbb{R}[x, (x+1)^{-1}]}{(x^2-1)} \\ &= \frac{\mathbb{R}[x, (x+1)^{-1}]}{((x+1)(x-1))} \\ &= \frac{\mathbb{R}[x, (x+1)^{-1}]}{(x-1)} \\ &\approx \mathbb{R} \end{aligned}$$

Here, since  $x+1$  has residue 2 in  $R = \mathbb{R}[x]/(x^2-1)$

Therefore, the required ring is  $\boxed{\frac{\mathbb{R}[x,y]}{(x^2-1, (x+1)y-1)} = \mathbb{R}}$

7. a

$$(2x) \text{ and } (2) \cap (x)$$

Let  $f(x) \in (2x)$ . Then there exists some polynomial  $g(x) \in \mathbb{Z}$  such that

$$f(x) = 2xg(x)$$

But this means that  $f(x) \in (2)$  (because  $xg(x)$  is a polynomial), and  $f(x) \in (x)$  (because  $2g(x)$  is a polynomial). Thus,  $f(x) \in (2) \cap (x)$ , and

$$(2x) \subseteq (2) \cap (x)$$

On the other hand, let  $p(x) \in (2) \cap (x)$ . Since  $p(x) \in (2)$ , there exists some polynomial  $h(x) \in \mathbb{Z}[x]$  such that

$$p(x) = 2h(x)$$

Furthermore,  $p(x) \in (x)$ , so

$$p(x) = xh_2(x)$$

So,  $2h(x) = xh_2(x)$ , for some  $h_2(x) \in \mathbb{Z}[x]$ . This means that  $h(0) = 0$ , so  $x$  divides  $h(x)$ ; that is,

$$h(x) = xq(x)$$

for some  $q(x) \in \mathbb{Z}[x]$ , and

$$p(x) = 2xq(x)$$

Thus,  $p(x) \in (2x)$ , and

$$(2) \cap (x) \subseteq (2x)$$

Finally,

$$(2) \cap (x) = (2x),$$

as required.

Isomorphism.

Define a map

$$f : \mathbb{Z}[x] \rightarrow \mathbb{F}_2[x] \times \mathbb{Z}, \quad f(p(x)) = (\bar{p}(x), p(0)),$$

where  $\bar{p}(x)$  is obtained from  $p(x)$  by modding all coefficients by 2. Moreover,  $f$  is clearly a homomorphism.

Now we find its kernel. Let  $p(x)$  be such that  $f(p(x)) = (0, 0)$ . Then

$$(\bar{p}(x), p(0)) = (0, 0)$$

Let

$$p(x) = \sum_{i=0}^n a_i x^i$$

Then  $p(0) = a_0$ , so we must have  $a_0 = 0$ . Moreover,

$$\bar{p}(x) = 0$$

means that  $\bar{a}_i = 0$ , where  $\bar{a}_i$  is obtained by modding  $a_i$  by 2. Thus,  $a_i$  must be even; finally,

$$p(x) = \sum_{i=1}^n 2b_i x^i = 2x \sum_{i=0}^{n-1} b_{i+1} x^i$$

for  $b_i$  such that  $a_i = 2b_i$ . Thus,  $p(x) \in (2x)$ , and

$$\ker f \subseteq (2x)$$

To prove the other inclusion, let  $g(x) \in (2x)$ . Then  $g(x) = 2xh(x)$ , for some polynomial  $h(x) \in \mathbb{Z}[x]$ . Thus,  $g(0) = 0$ , and  $\bar{g}(x) = 0$ , since all coefficients of  $g(x)$  are even. Finally,  $f(g(x)) = 0$ , so  $g(x) \in \ker f$  and

$$(2x) \subseteq \ker f$$

Finally,

$$\ker f = (2x)$$

Now we want to find the image of  $f$ . First of all, if

$$f(p(x)) = (\bar{p}(x), n),$$

then  $p(0) = n$ . Now let

$$p(x) = \sum_{i=0}^m a_i x^i$$

Then  $p(0) = a_0 = n$ , and

$$\bar{p}(x) = \sum_{i=0}^m \bar{a}_i x^i$$

This means that

$$\bar{p}(0) = \bar{a}_0 = \bar{n}$$

This means precisely that  $p(0) \equiv n$  modulo 2. Thus, the image of  $f$  is contained inside the set described in the exercise. On the other hand, let  $(\bar{p}(x), n)$  be such that  $\bar{p}(0) \equiv n$  modulo 2. Let

$$\bar{p}(x) = \sum_{i=0}^n c_i x^i,$$

where  $c_i \in \mathbb{F}_2$ , and  $c_0 \equiv n$  modulo 2 (this must hold since  $\bar{p}(0) = c_0$ ). Now define

$$p(x) = \sum_{i=0}^n a_i x^i,$$

where

$$a_i = \begin{cases} n & \text{if } i = 0 \\ 0 & \text{if } c_i = 0, i \geq 1 \\ 1 & \text{if } c_i = 1, i \geq 1 \end{cases}$$

Now clearly

$$f(p(x)) = (\bar{p}(x), n),$$

so the set described in the exercise is contained in the image of  $f$ .

Now we use the First Isomorphism Theorem to conclude that

$$\mathbb{Z}[x]/(2x) \approx S,$$

where  $S$  is the set described in the exercise.

## Result

To prove that  $(2) \cap (x) = (2x)$ , show that the two inclusions hold.

For isomorphism, use the First Isomorphism Theorem.

8. a

(a)

$$\underline{IJ \subseteq I \cap J}$$

Let  $r \in IJ$ . Then

$$r = \sum_{i=1}^n x_i y_i$$

for some  $x_i \in I, y_i \in J, n \in \mathbb{N}$ . Since  $x_i \in I$  and  $I$  is an ideal, we conclude that  $x_i y_i \in I$ . Since  $I$  is closed under addition, we conclude that  $r \in I$ .

Similarly, since  $y_i \in J$  and  $J$  is an ideal, we conclude that  $x_i y_i \in J$ . Since  $J$  is closed under addition, we conclude that  $r \in J$ .

Thus,  $r \in I \cap J$ , so

$$IJ \subseteq I \cap J$$

$$\underline{I \cap J \subseteq IJ}$$

Since  $I + J = R$ , specially we can find  $a \in I, b \in J$  such that  $a + b = 1$ . Now let  $r \in I \cap J$ . Then

$$r = r \cdot 1 = r(a + b) = ra + rb = ar + rb$$

Since  $r \in I \cap J, a \in I, b \in J$ , the above equality means that  $r \in IJ$ . Thus

$$I \cap J \subseteq IJ$$

#### Conclusion.

Finally, we conclude that

$$IJ = I \cap J$$

(b)

We observe  $b - a$ . Since  $b - a \in R$  and  $I + J = R$ , there exist some  $m \in I, n \in J$  such that

$$b - a = m + n$$

Now notice that  $m + n = m - (-n)$ , so we can set  $n' = -n$  to get

$$b - a = m - n'$$

From this, we also get

$$a + m = b + n'$$

Thus, if we define  $x$  as  $x = a + m = b + n'$ , we get:

$$x - a = m \in I$$

$$x - b = n' \in J$$

Thus, we have proven the statement.

(c)

We define a map

$$f : R \rightarrow (R/I) \times (R/J), \quad f(r) = (r + I, r + J)$$

To prove that  $f$  is a homomorphism, let  $x, y \in R$ . Then

$$\begin{aligned} f(x + y) &= ((x + y) + I, (x + y) + J) \\ &= ((x + I) + (y + I), (x + J) + (y + J)) \\ &= (x + I, x + J) + (y + I, y + J) \\ &= f(x) + f(y) \end{aligned}$$

$$\begin{aligned} f(xy) &= ((xy) + I, (xy) + J) \\ &= ((x + I)(y + I), (x + J)(y + J)) \\ &= (x + I, x + J)(y + I, y + J) \\ &= f(x)f(y) \end{aligned}$$

Also,  $1 + I$  is the multiplicative identity in  $R/I$  and  $1 + J$  is the multiplicative in  $R/J$ , and

$$f(1) = (1 + I, 1 + J)$$

So,  $f$  is truly a homomorphism.

Now we want to prove that it is also a bijection. We find its kernel. Let  $r \in R$  be such that

$$f(r) = (I, J)$$

(we put  $(I, J)$  on the RHS because  $(I, J)$  is the additive inverse in  $(R/I) \times (R/J)$ ), then

$$(r + I, r + J) = (I, J)$$

This means that  $r + I = I$  and  $r + J = J$ . It follows that  $r \in I$  and  $r \in J$ , respectively. Thus,  $r \in I \cap J$ . Furthermore,

$$\ker f \subseteq I \cap J$$

However,  $I \cap J = IJ = 0$ , so  $\ker f = 0$ , which means that the kernel is trivial and  $f$  is injective.

Now we want to prove that  $f$  is also surjective. Let  $(a + I, b + J) \in (R/I) \times (R/J)$ . By (b), there exists  $x \in R$  such that  $x - a \in I$ ,  $x - b \in J$ . But this means precisely that  $x + I = a + I$  and  $x + J = b + J$ . Thus,

$$f(x) = (x + I, x + J) = (a + I, b + J),$$

so  $f$  is surjective.

Thus,  $f$  is bijective. Since it is also a homomorphism, we conclude that  $f$  is an isomorphism; that is,

$$R \approx (R/I) \times (R/J),$$

as required.

(d)

Let  $(e + I, e' + J)$  be some idempotent element in  $(R/I) \times (R/J)$ . Then

$$(e + I, e' + J) = (e + I, e' + J)^2 = (e^2 + I, e'^2 + J)$$

Thus,

$$\begin{aligned} e^2 + I = e + I &\implies e^2 - e \in I \\ e'^2 + J = e' + J &\implies e'^2 - e' \in J \end{aligned}$$

HINTS:

(a)  $IJ \subseteq I \cap J$  follows almost immediately from the definition of  $IJ$ . On the other hand,  $1 = a + b$  for some  $a \in I, b \in J$ , and

$$x = x \cdot 1 = ax + xb$$

(b) Observe  $b - a$ .

(c) Observe a map  $f : R \rightarrow (R/I) \times (R/J), f(r) = (r + I, r + J)$ .

(d) Such elements  $(e + I, e' + J)$ , where  $e^2 - e \in I, e'^2 - e' \in J$ .

## Section 7

1. a

Notice that we only need to prove that each nonzero element of  $R$  is invertible, since all other axioms of a field are satisfied.

Let  $x \in R$  be some nonzero element. Define a function

$$f : R \rightarrow R, \quad f(y) = xy$$

(So, we just multiply  $y$  by  $x$ . Also, it is worth mentioning that this is **not** a homomorphism.)

We will prove that  $f$  is injective. Let  $y_1, y_2 \in R$  be such that  $f(y_1) = f(y_2)$ . Then

$$xy_1 = xy_2 \implies x(y_1 - y_2) = 0$$

Since  $R$  is a domain and  $x \neq 0$ , we conclude that  $y_1 - y_2 = 0$ ; that is,  $y_1 = y_2$ .

Thus,  $f$  is an injective function from a finite set to itself. This means that  $f$  is also a bijection, so there exists some  $z \in R$  such that  $f(z) = 1$ . Thus,

$$xz = 1,$$

which means that  $x$  is invertible! (With  $z = x^{-1}$ .)

Since  $x$  was arbitrarily taken nonzero element of  $R$ , it follows that all of them are invertible. Thus,  $R$  is a field.

### Result

It is sufficient to prove that each nonzero element of  $R$  is invertible. (Why?)

Fix some  $x \in R, x \neq 0$ , and define

$$f : R \rightarrow R, \quad f(y) = xy$$

What can you tell about  $f$ ?

2. a



Suppose that  $R[x]$  is not a domain. Let  $p(x), q(x) \in R[x]$  be nonzero polynomials that such  $p(x)q(x) = 0$ . Let

$$p(x) = a_0 + a_1x + \dots + a_nx^n$$

and

$$q(x) = b_0 + b_1x + \dots + b_mx^m,$$

where  $b_m \neq 0$  and  $a_n \neq 0$ . Then

$$p(x)q(x) = (a_0b_0) + (a_1b_0 + a_0b_1)x + \dots + (a_nb_m)x^{n+m}$$

Since  $p(x)q(x) = 0$ , we must have that  $a_nb_m = 0$ . However this is impossible since  $R$  is a domain,  $a_n \neq 0$ , and  $b_m \neq 0$ .

Thus,  $R[x]$  must be a domain.

## Result

2 of 2

Suppose that  $p(x)q(x) = 0$  for some  $p(x) \neq 0, q(x) \neq 0$ . Analyze the highest nonzero coefficients of  $p(x)$  and  $q(x)$ .

### 3. a

Take any ring  $R$  with 15 element. We begin by inspecting the additive abelian group  $R^+ = (R, +)$ .

Since  $15 = 3 \cdot 5$ , by the First Sylow Theorem, we conclude that there exists a  $p$ -subgroup  $H$  of  $R^+$  of order 3. By the Third Sylow Theorem, the number of such subgroups is  $3k + 1$ , with  $k$  a nonnegative integer, and  $3k + 1$  divides  $|R^+| = 15$ . So, there is only one such subgroup. By the Corollary of the Second Sylow Theorem, we now conclude that  $H$  is normal in  $R^+$ .

Similarly, there exists a  $p$ -subgroup  $K$  of  $R^+$  of order 5. By the Third Sylow Theorem, the number of such subgroups is  $5k + 1$ , with  $k$  a nonnegative integer, and  $5k + 1$  divides  $|R^+| = 15$ . So, there is only one such subgroup. By the Corollary of the Second Sylow Theorem, we now conclude that  $K$  is normal in  $R^+$ .

Now let  $x \in H \cap K$ . Since  $H$  is of order 3, we must have  $3x = 0$ . Since  $K$  is of order 5, we must have  $5x = 0$ . (These equalities are written additively!) Furthermore, 3 and 5 are relatively prime, so there exist integers  $m, n$  such that  $1 = 3m + 5n$ . Thus,

$$x = 1x = (3m + 5n)x = m(3x) + n(5x) = 0$$

So,  $H \cap K = \{0\}$ . This means that

$$H \times K \approx HK$$

Furthermore,  $|H \times K| = 15$ , so  $|HK| = 15$ , and we must have that  $HK = R^+$  (since  $HK \subseteq R^+$  and they have the same number of elements). Thus,

$$R^+ \approx H \times K$$

Since  $H$  is of order 3, which is a prime number, there exists some  $a \in H$  such that  $H = \{0, a, 2a\}$  (in other words,  $H$  is cyclic). Now the mapping  $na \rightsquigarrow n$  forms an isomorphism  $H \rightarrow \mathbb{Z}/3\mathbb{Z}$ . Thus,

$$H \approx \mathbb{Z}/3\mathbb{Z}$$

Similarly,

$$K \approx \mathbb{Z}/5\mathbb{Z}$$

Therefore,

$$R^+ \approx \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$$

Furthermore, since 3 and 5 are relatively prime,

$$\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \approx \mathbb{Z}/15\mathbb{Z}$$

Thus,

$$R^+ \approx \mathbb{Z}/15\mathbb{Z}$$

Since  $\mathbb{Z}/15\mathbb{Z}$  is cyclic,  $R^+$  must also be cyclic, so

$$R^+ = \{\bar{0}, \bar{1}, \bar{2}, \dots, \bar{14}\},$$

where  $\bar{x} = \underbrace{\bar{1} + \dots + \bar{1}}_{x \text{ times}}$ .

Since  $R^+$  and  $R$  are under the same set,

$$R = \{\bar{0}, \bar{1}, \bar{2}, \dots, \bar{14}\}$$

We will prove that the multiplicative identity can generate  $R^+$ . It suffices to show that  $n\bar{1} \neq \bar{0}$ , for all  $n < 15$ .

Suppose that  $n\bar{1} = \bar{0}$ , for  $n < 15$ . Then  $nr = n\bar{1}r = \bar{0}$ , for all  $r \in R$ , so none of them can generate  $R^+$ .

Thus, we can take that  $\bar{1}$  is the multiplicative identity in  $R$ .

Now it is clear that  $R$ , as a ring, is isomorphic to the ring  $(\mathbb{Z}/15\mathbb{Z}, +, \cdot)$ . (The mapping  $R \rightarrow \mathbb{Z}/15\mathbb{Z}$ ,  $\bar{x} \mapsto x$  is an isomorphism.) However, in  $\mathbb{Z}/15\mathbb{Z}$  we have that  $3 \cdot 5 = 0$ , and  $3 \neq 0$ ,  $5 \neq 0$ , so  $\mathbb{Z}/15\mathbb{Z}$  is not a domain. Since  $R$  is isomorphic to it, it also cannot be a domain.

## Result

3 of 3

No such domains exist. HINT: Prove that all ring of order 15 are isomorphic to  $\mathbb{Z}/15\mathbb{Z}$  and that it is not an integral domain.

## 4. a

What we really need to prove is that for every  $f(x), g(x) \in F[[x]]$ ,  $g(x) \neq 0$  there exist some  $h(x) \in F[[x]]$  and  $n$  nonnegative integer such that

$$f(x)/g(x) = h(x)/x^n$$

Since  $g(x) \neq 0$ ,

$$g(x) = \sum_{i=k}^{\infty} a_i x^i,$$

where  $a_k \neq 0$ , for some  $k$  nonnegative integer (possibly zero). By Exercise 11.2.2, we conclude that

$$\tilde{g}(x) = a_k + a_{k+1}x + \dots$$

is invertible. So there exists some  $\tilde{g}^{-1}(x) \in F[[x]]$  such that

$$\tilde{g}(x)\tilde{g}^{-1}(x) = 1$$

Now we first see that

$$g(x) = x^k \tilde{g}(x),$$

so

$$f(x)/g(x) = f(x)/(x^k \tilde{g}(x))$$

Now we want to "multiply the numerator and denominator by  $\tilde{g}^{-1}(x)$ ". Since clearly  $\tilde{g}^{-1}(x) \neq 0$  (otherwise,  $\tilde{g}(x)\tilde{g}^{-1}(x) = 1$  would be impossible), the fraction

$$(f(x)\tilde{g}^{-1}(x))/(x^k \tilde{g}(x)\tilde{g}^{-1}(x))$$

makes sense. Furthermore, by cross-multiplying we check that

$$f(x)/(x^k \tilde{g}(x)) = (f(x)\tilde{g}^{-1}(x))/(x^k \tilde{g}(x)\tilde{g}^{-1}(x))$$

Now all that is left to conclude is that

$$(f(x)\tilde{g}^{-1}(x))/(x^k \tilde{g}(x)\tilde{g}^{-1}(x)) = (f(x)\tilde{g}^{-1}(x))/x^k$$

Thus, if we set  $h(x) = f(x)\tilde{g}^{-1}(x)$ , we have that

$$f(x)/g(x) = h(x)/x^k,$$

as required.

Notice that if we formally divide a series  $h(x)$  by  $x^k$ , we get an expression of the form

$$\sum_{i=-m}^{\infty} a_i x^i$$

for some  $m$ .

## Result

2 of 4

We need to check that, for all  $f(x), g(x) \in F[[x]]$ ,  $g(x) \neq 0$ , there exist some  $h(x) \in F[[x]]$  and  $n$  nonnegative integer such that

$$f(x)/g(x) = h(x)/x^n$$

Furthermore, one easily sees that, after formally dividing  $h(x)$  by  $x^n$ , we get expression of the form

$$\sum_{i=-m}^{\infty} a_i x^i$$

for some  $m$ .

## 5. a

Denote this set by  $\mathcal{R}$ . We check properties from the definition of a ring.

$(\mathcal{R}, +)$  is a commutative group.

Closure.

Let  $a_1/b_1, a_2/b_2 \in \mathcal{R}$ . Then

$$a_1/b_1 + a_2/b_2 = (a_1 b_2 + a_2 b_1)/b_1 b_2$$

Since  $b_1, b_2 \in S$ , we have that  $b_1 \neq 0, b_2 \neq 0$ . Since  $R$  is a domain,  $b_1 b_2 \neq 0$ . Thus,  $b_1, b_2 \in S$ , which means that  $a_1/b_1 + a_2/b_2 \in \mathcal{R}$ .

Associativity.

Let  $a_1/b_1, a_2/b_2, a_3/b_3 \in \mathcal{R}$ . Then

$$\begin{aligned} (a_1/b_1 + a_2/b_2) + a_3/b_3 &= (a_1 b_2 + a_2 b_1)/b_1 b_2 + a_3/b_3 \\ &= ((a_1 b_2 + a_2 b_1) b_3 + a_3 b_1 b_2)/((b_1 b_2) b_3) \\ &= (a_1 b_2 b_3 + a_2 b_1 b_3 + a_3 b_1 b_2)/(b_1 b_2 b_3) \end{aligned}$$

(Notice that we used that multiplying in  $R$  is associative and commutative in the last equality. We also used the distributive law.)

On the other hand,

$$\begin{aligned} a_1/b_1 + (a_2/b_2 + a_3/b_3) &= a_1/b_1 + (a_2b_3 + a_3b_2)/b_2b_3 \\ &= (a_1(b_2b_3) + (a_2b_3 + a_3b_2)b_1)/(b_1(b_2b_3)) \\ &= (a_1b_2b_3 + a_2b_1b_3 + a_3b_1b_2)/(b_1b_2b_3) \end{aligned}$$

Thus,

$$(a_1/b_1 + a_2/b_2) + a_3/b_3 = a_1/b_1 + (a_2/b_2 + a_3/b_3)$$

which we needed to show.

#### Identity.

We will prove that  $0/1$  is the additive identity. First,  $0/1 \in \mathcal{R}$  since  $1 \neq 0$ , so  $1 \in S$ . Let  $a/b \in \mathcal{R}$ . Then

$$a/b + 0/1 = (a \cdot 1 + 0 \cdot b)/(b \cdot 1) = a/b$$

#### Inverse.

Let  $a/b \in \mathcal{R}$ . Then  $(-a)/b \in \mathcal{R}$ , and

$$a/b + (-a)/b = (ab - ab)/b^2 = 0/b^2 = 0/1$$

(the last equality is checked by cross-multiplying, which yields  $0 = 0$ ). Thus, each element of  $\mathcal{R}$  has an additive inverse.

#### Commutativity.

Let  $a_1/b_1, a_2/b_2 \in \mathcal{R}$ . Then

$$a_1/b_2 + a_2/b_2 = (a_1b_2 + a_2b_1)/b_1b_2 \stackrel{(1)}{=} (a_2b_1 + a_1b_2)/b_2b_1 = a_2/b_2 + a_1/b_1,$$

where in (1) we used the commutativity of addition and multiplication in  $R$ .

**$(\mathcal{R}, \cdot)$  is associative, commutative, and has an identity.**

#### Closure.

Let  $a_1/b_1, a_2/b_2 \in \mathcal{R}$ . Then

$$a_1/b_1 \cdot a_2/b_2 = (a_1b_2 + a_2b_1)/b_1b_2$$

Since  $b_1, b_2 \in S$ , we have that  $b_1 \neq 0, b_2 \neq 0$ . Since  $R$  is a domain,  $b_1b_2 \neq 0$ . Thus,  $b_1, b_2 \in S$ , which means that  $a_1/b_1 \cdot a_2/b_2 \in \mathcal{R}$ .

#### Associativity.

Let  $a_1/b_1, a_2/b_2, a_3/b_3 \in \mathcal{R}$ . Then

$$\begin{aligned} (a_1/b_1 \cdot a_2/b_2) \cdot a_3/b_3 &= a_1a_2/b_1b_2 \cdot a_3/b_3 \\ &= ((a_1a_2)a_3)/((b_1b_2)b_3) \\ &= (a_1a_2a_3)/(b_1b_2b_3) \end{aligned}$$

(Notice that we used that multiplying in  $R$  is associative in the last equality.)

On the other hand,

$$\begin{aligned} a_1/b_1 \cdot (a_2/b_2 \cdot a_3/b_3) &= a_1/b_1 \cdot a_2a_3/b_2b_3 \\ &= (a_1(a_2a_3))/(b_1(b_2b_3)) \\ &= (a_1a_2a_3)/(b_1b_2b_3) \end{aligned}$$

Thus,

$$(a_1/b_1 \cdot a_2/b_2) \cdot a_3/b_3 = a_1/b_1 \cdot (a_2/b_2 \cdot a_3/b_3)$$

which we needed to show.

#### Identity.

We will prove that  $1/1$  is the additive identity. First,  $1/1 \in \mathcal{R}$  since  $1 \neq 0$ , so  $1 \in S$ . Let  $a/b \in \mathcal{R}$ . Then

$$a/b \cdot 1/1 = (a \cdot 1)/(b \cdot 1) = a/b$$

#### Commutativity.

Let  $a_1/b_1, a_2/b_2 \in \mathcal{R}$ . Then

$$a_1/b_2 \cdot a_2/b_2 = a_1 a_2 / b_1 b_2 \stackrel{(1)}{=} a_2 a_1 / b_2 b_1 = a_2/b_2 \cdot a_1/b_1,$$

where in (1) we used the commutativity of multiplication in  $R$ .

#### Distributive law.

Let  $a_1/b_1, a_2/b_2, a_3/b_3 \in \mathcal{R}$ . Then

$$\begin{aligned} (a_1/b_1 + a_2/b_2) \cdot a_3/b_3 &= (a_1 b_2 + a_2 b_1) / (b_1 b_2) \cdot a_3/b_3 \\ &= ((a_1 b_2 + a_2 b_1) a_3) / ((b_1 b_2) b_3) \\ &= (a_1 a_3 b_2 + a_2 a_3 b_1) / (b_1 b_2 b_3) \end{aligned}$$

(We used the associativity and commutativity of multiplication, and the distributive law in the last equality.)

On the other hand,

$$\begin{aligned} a_1/b_1 \cdot a_3/b_3 + a_2/b_2 \cdot a_3/b_3 &= (a_1 a_3) / (b_1 b_3) + (a_2 a_3) / (b_2 b_3) \\ &= (a_1 a_3 b_2 b_3 + a_2 a_3 b_1 b_3) / (b_1 b_2 b_3^2) \\ &= (a_1 a_3 b_2 + a_2 a_3 b_1) / (b_1 b_2 b_3) \end{aligned}$$

(The last equality is checked by cross-multiplying.)

Thus, even the distributive law holds.

### **Result**

4 of 4

Check properties from the definition of a ring.

## Section 8

1. a



We prove that there are no principal ideal is also a maximal ideal. We first rule out ideals generated by a constant polynomial, and then by other polynomials.

#### Constant polynomials.

Let  $n \in \mathbb{Z}$ , and observe the ideal  $(n)$ . If  $n = \pm 1$ , then clearly  $(n) = \mathbb{Z}[x]$ , so  $(n)$  is not maximal.

If  $n \neq \pm 1$ , then  $x \notin (n)$ . To prove this, suppose that  $x \in (n)$ . Then there exists a polynomial  $f(x) \in \mathbb{Z}[x]$  such that

$$x = nf(x)$$

The degrees of polynomials on both sides must match, so  $\deg f(x) = 1$ , and  $f(x) = a_0 + a_1x$  for some  $a_0, a_1 \in \mathbb{Z}$ . Now the equality  $x = nf(x)$  becomes

$$x = na_0 + na_1x$$

So,

$$na_1 = 1$$

However, this is impossible since  $n \neq \pm 1$ , so it is not a unit in  $\mathbb{Z}$  (the above equality actually tells us that  $a_1 = n^{-1}$ ). Thus,  $x \notin (n)$ .

Now define  $I = (n, x)$ . We will prove that  $1 \notin I$ . Suppose that  $1 \in I$ . Then we have that there exist polynomials  $f(x), g(x) \in \mathbb{Z}[x]$  such that

$$1 = nf(x) + xg(x)$$

The degrees of both sides must match, so  $f(x)$  is of degree 0 (constant polynomial), and we must have  $g(x) = 0$ . Thus,  $f(x) = a$ , and

$$na = 1$$

Similarly to before, this is impossible since  $n$  is not a unit in  $\mathbb{Z}$ .

Thus,  $1 \notin I$ , so  $I \neq \mathbb{Z}[x]$ .

Now we have that

$$(n) \subset (n, x) \subset \mathbb{Z}[x],$$

so  $(n)$  is not maximal.

#### Nonconstant polynomials.

Now observe  $(f(x))$ , for some nonconstant polynomial  $f(x) \in \mathbb{Z}[x]$ . First of all,

$$(f(x)) = \{f(x)g(x) \mid g(x) \in \mathbb{Z}[x]\},$$

so there are no nonzero constant polynomials in  $(f(x))$ . Now let  $p \in \mathbb{Z}$  be a prime integer which does not divide the leading coefficient of  $f(x)$ . We will prove that

$$(f(x)) \subset (p, f(x)) \subset \mathbb{Z}[x]$$

The first strict inclusion follows from the fact that  $g(x) = p$  is a constant polynomial, so it cannot be contained in  $(f(x))$ .

To prove that the other inclusion is also strict, suppose that  $1 \in (p, f(x))$ . Then there exist polynomials  $g(x), h(x)$  such that

$$1 = pg(x) + f(x)h(x)$$

We write this equation as

$$pg(x) = 1 - f(x)h(x) \tag{1}$$



Now suppose that

$$\begin{aligned} f(x) &= a_0 + a_1x + \dots + a_nx^n \\ h(x) &= c_0 + c_1x + \dots + c_mx^m \end{aligned}$$

Thus,

$$f(x)h(x) = a_0c_0 + (a_0c_1 + a_1c_0)x + \dots + (a_{n-1}c_m + a_nc_{m-1})x^{n+m-1} + a_nc_mx^{n+m}$$

Since in (1) both sides are of the same degree, we get that

$$g(x) = b_0 + b_1x + \dots + b_{n+m}x^{n+m}$$

Now we write (1) as

$$\begin{aligned} a_nc_m &= pb_{n+m} \\ a_{n-1}c_m + a_nc_{m-1} &= pb_{n+m-1} \\ &\vdots \\ a_0c_1 + a_1c_0 &= pb_1 \\ a_0c_0 - 1 &= pb_0 \end{aligned}$$

From the first equality we have that  $p$  divides  $a_nc_m$ . Since  $p$  is prime, it divides either  $a_n$  or  $c_m$ . Since it does not divide  $a_n$  (the leading coefficient of  $f(x)$ ), it divides  $c_m$ .

From the second equation, we now get that

$$a_nc_{m-1} = pb_{n+m-1} - a_{n-1}c_m$$

Since  $p$  divides the right side, it must also divide the left side. Thus,  $p$  divides  $a_nc_{m-1}$ . Since it does not divide  $a_n$ , it divides  $c_{m-1}$ .

Inductively, we now see that we can conclude that  $p$  divides all coefficients of  $g(x)$ . Now look at the last equation of the system:

$$1 = a_0c_0 - pb_0$$

Since  $p$  divides the right side, we get that  $p$  divides 1, which is impossible.

Thus,  $1 \notin (p, f(x))$ , so  $(p, f(x)) \neq \mathbb{Z}[x]$ .

To conclude, we now have that

$$(f(x)) \subset (p, f(x)) \subset \mathbb{Z}[x],$$

so  $(f(x))$  is not a maximal ideal.

## Result

There are no maximal ideals in  $\mathbb{Z}[x]$  which are also principal ideals.

2. a

(a)

We first observe

$$I_1 = ((1, 0)) = \{(x, 0) \mid x \in \mathbb{R}\}$$

This is obviously an ideal in  $\mathbb{R} \times \mathbb{R}$ . Now suppose that  $I_1 \subset I$  for some ideal  $I$ . Then  $I$  must contain an element of the form  $(a, b)$ , where  $b \neq 0$ . If  $a = 0$ , then  $(a + 1, b) \in I$ , since  $(1, 0) \in I_1 \subset I$  and  $(a, b) \in I$ . Thus,  $I$  contains an element of the form  $(a', b')$  with  $a' \neq 0, b' \neq 0$ . Furthermore, it must now contain  $(1, 1)$ , since  $(1, 1) = (a, b)(a^{-1}, b^{-1})$ . But now  $I = \mathbb{R} \times \mathbb{R}$ , since  $(1, 1)$  is the multiplicative identity in  $\mathbb{R} \times \mathbb{R}$ !

Thus,  $I_1$  is a maximal ideal.

Similarly we can show that

$$I_2 = ((0, 1))$$

is a maximal ideal.

Now it is clear that the zero ideal is not maximal.

Take any  $I$  in  $\mathbb{R} \times \mathbb{R}$ . If it is the zero ideal, then it is not maximal. If it contains an element  $(a, 0)$ ,  $a \neq 0$ , then it clearly contains the entire  $I_1$ , so it is either  $I_1$  or  $\mathbb{R} \times \mathbb{R}$  (because of the maximality of  $I_1$ ). If it contains an element  $(0, b)$ ,  $b \neq 0$ , then it is either  $I_2$  or  $\mathbb{R} \times \mathbb{R}$ .

Thus,  $I_1$  and  $I_2$  are the only maximal ideals in  $\mathbb{R} \times \mathbb{R}$ .

(b)

Let  $f(x) \in \mathbb{R}[x]$ . Then we can divide  $f(x)$  by  $x^2$  to get

$$f(x) = g(x)x^2 + (ax + b)$$

for some polynomial  $g(x) \in \mathbb{R}[x]$  and  $a, b \in \mathbb{R}$ . This means that

$$f(x) - (ax + b) = g(x)x^2 \in (x^2)$$

Thus,

$$f(x) + (x^2) = (ax + b) + (x^2)$$

Now we get that

$$\mathbb{R}[x]/(x^2) = \{ax + b + (x^2) \mid a, b \in \mathbb{R}\}$$

Let  $I$  be some maximal ideal in  $\mathbb{R}[x]/(x^2)$ . If  $I$  contains  $c + (x^2)$ , for some  $c \neq 0$ , then  $I$  contains  $1 + (x^2)$ , because  $1 + (x^2) = (c + (x^2))(c^{-1} + (x^2))$ . Furthermore,  $1 + (x^2)$  is a multiplicative identity in  $\mathbb{R}[x]/(x^2)$ , so  $I = \mathbb{R}[x]/(x^2)$ . Thus,  $I$  cannot be maximal.

If  $I$  contains some  $(ax + b) + (x^2)$ , with  $a \neq 0, b \neq 0$ , then again  $I = \mathbb{R}[x]/(x^2)$ . First of all,  $((ax + b) + (x^2))(x + (x^2)) = (ax^2 + bx) + (x^2) = bx + (x^2)$ . Thus,  $bx + (x^2) \in I$ .

This also means that  $ax + (x^2) \in I$ , since  $(bx + (x^2))(ab^{-1} + (x^2)) = ax + (x^2)$ .

But now  $b + (x^2) = (ax + b) + (x^2) - (ax + (x^2))$ , so  $b + (x^2) \in I$ . From before we now know that  $I = \mathbb{R}[x]/(x^2)$ .

Notice that we ruled out all ideals except  $I = (x + (x^2))$ . First of all,

$$I = \{(x + (x^2))(ax + b + (x^2)) \mid a, b \in \mathbb{R}\} = \{bx + (x^2) \mid b \in \mathbb{R}\}$$

Thus,  $I \neq \mathbb{R}[x]/(x^2)$ . Now suppose that  $J$  is some ideal such that  $I \subset J$ . Then  $J$  must contain some element  $ax + b + (x^2)$ , with  $a, b \in \mathbb{R}, b \neq 0$ . But from before we know that  $J = \mathbb{R}[x]/(x^2)$ .

Thus, the only maximal ideal is  $(x + (x^2))$ .

(c)

Define a mapping

$$\varphi : \mathbb{R}[x] \rightarrow \mathbb{R} \times \mathbb{R}, \quad \varphi(f(x)) = (f(1), f(2))$$

This is clearly a homomorphism. Now let  $f(x) \in \ker \varphi$ . Then  $f(1) = f(2) = 0$ , so  $x - 1$  and  $x - 2$  divide  $f(x)$ . This means that

$$f(x) = (x - 1)(x - 2)g(x) = (x^2 - 3x + 2)g(x)$$

for some polynomial  $g(x) \in \mathbb{R}[x]$ . Thus,  $\ker \varphi \subseteq (x^2 - 3x + 2)$ . Furthermore,  $(x^2 - 3x + 2) \subseteq \ker \varphi$  is trivial. Thus,

$$\ker \varphi = (x^2 - 3x + 2)$$

Now to prove that  $\varphi$  is surjective. Let  $(a, b) \in \mathbb{R} \times \mathbb{R}$ . Then

$$\varphi(f(x)) = (a, b)$$

yields the equation

$$(f(1), f(2)) = (a, b),$$

which yields the system

$$\begin{aligned} f(1) &= a \\ f(2) &= b \end{aligned}$$

So, we need to unknowns. We set  $f(x) = c_0 + c_1x$ . Then the system becomes

$$\begin{aligned} c_0 + c_1 &= a \\ c_0 + 2c_1 &= b \end{aligned}$$

The solution is  $c_1 = b - a$ ,  $c_0 = 2a - b$ . Thus, the solution exists, so  $\varphi(f(x)) = (a, b)$ , and  $\varphi$  is surjective.

By the First Isomorphism Theorem,

$$\mathbb{R}[x]/(x^2 - 3x + 2) \approx \mathbb{R} \times \mathbb{R}$$

By (a), the only maximal ideals in  $\mathbb{R} \times \mathbb{R}$  are

$$I_1 = \{(a, 0) \mid a \in \mathbb{R}\} \quad \text{and} \quad I_2 = \{(0, b) \mid b \in \mathbb{R}\}$$

Now notice that  $\varphi(f(x)) \in I_1$  if and only if  $f(2) = 0$ , which is if and only if  $f(x) \in (x - 2)$ . Thus,  $\varphi^{-1}(I_1) = (x - 2)$ . This also means that  $\mathbb{R}[x]/(x - 2)$  is a maximal ideal in  $\mathbb{R}[x]/(x^2 - 3x + 2)$ .

Similarly,  $\varphi^{-1}(I_2) = (x - 1)$ , so the other maximal ideal is  $\mathbb{R}[x]/(x - 1)$ .

(d)

We will prove that  $x^2 + x + 1$  is irreducible. Suppose that it is not. Then there exist nonconstant polynomials  $f(x), g(x) \in \mathbb{R}[x]$  such that

$$x^2 + x + 1 = f(x)g(x)$$

Since degrees must be equal, we conclude that  $f(x)$  and  $g(x)$  are of degree 1. Thus,  $f(x) = a_0 + a_1x$  and  $g(x) = b_0 + b_1x$ . The above equality now becomes

$$x^2 + x + 1 = (a_0 + a_1x)(b_0 + b_1x) = (a_1b_1)x^2 + (a_0b_1 + a_1b_0)x + a_0b_0$$

From this,

$$a_0b_0 = 1 \implies a_0 = \frac{1}{b_0}$$

and

$$a_1b_1 = 1 \implies a_1 = \frac{1}{b_1}$$

Furthermore,

$$a_0b_1 + a_1b_0 = 1 \implies \frac{b_1}{b_0} + \frac{b_0}{b_1} = 1$$

Multiplying by  $b_0b_1$ ,

$$2b_0b_1 = b_0b_1 \implies b_0b_1 = 0$$

Thus, either  $b_0 = 0$ , or  $b_1 = 0$ . However, both are impossible, since  $a_0b_0 = 1$  and  $a_1b_1 = 1$ .

Thus,  $x^2 + x + 1$  is irreducible, so, by Proposition 11.8.4 (a),  $(x^2 + x + 1)$  is a maximal ideal in  $\mathbb{R}[x]$ , hence by Proposition 11.8.2 (b)  $\mathbb{R}[x]/(x^2 + x + 1)$  is a field. Now we know that the only ideals in  $\mathbb{R}[x]/(x^2 + x + 1)$  are the zero ideal  $(0)$  and  $\mathbb{R}[x]/(x^2 + x + 1)$  itself. Thus, the zero ideal  $(0)$  is the only maximal ideal.

---

## Result

5 of 5

(a)  $((1, 0))$  and  $((0, 1))$ .

(b)  $(x + (x^2))$

(c)  $\mathbb{R}[x]/(x - 2)$  and  $\mathbb{R}[x]/(x - 1)$ .

(d) Only the zero ideal  $(0)$ .

3. a

### First case

By Proposition 11.8.2 (b), the statement is equivalent to the statement that  $(x^3 + x + 1)$  is a maximal ideal in  $\mathbb{F}_2[x]$ . By Proposition 11.8.4 (a), this is equivalent to  $x^3 + x + 1$  being irreducible in  $\mathbb{F}_2[x]$ .

Suppose that

$$x^3 + x + 1 = f(x)g(x),$$

for some nonconstant polynomials  $f(x), g(x) \in \mathbb{F}_2[x]$ . Thus, both  $f(x)$  and  $g(x)$  are of degree 1 or more. Furthermore,  $f(x)g(x)$  must be of degree 3, so it means that one of them is of degree 2, while the other is of degree 1. Without loss of generality we assume that  $\deg f(x) = 2, \deg g(x) = 1$ . Thus,

$$f(x) = a_0 + a_1x + a_2x^2$$

$$g(x) = b_0 + b_1x$$

where  $a_i, b_i \in \mathbb{F}_2$ . Since  $a_2, b_1 \neq 0$ , we conclude that  $a_2 = b_1 = 1$ . Thus,

$$f(x) = a_0 + a_1x + x^2$$

$$g(x) = b_0 + x$$

Now the equation

$$x^3 + x + 1 = f(x)g(x)$$

becomes

$$x^3 + x + 1 = (a_0 + a_1x + x^2)(b_0 + x),$$

or, with full details,

$$x^3 + x + 1 = x^3 + (b_0 + a_1)x^2 + (a_1b_0 + a_0)x + a_0b_0$$

This yields the system in  $\mathbb{F}_2$ :

$$a_0b_0 = 1$$

$$a_1b_0 + a_0 = 1$$

$$a_1 + b_0 = 0$$

From the first equation we get that  $a_0 = b_0 = 1$  (if either of them were 0, we get  $a_0b_0 = 0$ ).

Plugging this into the second equation,

$$a_1 + 1 = 1 \implies a_1 = 0$$

Plugging  $a_1 = 0, b_0 = 1$  into the third:

$$1 \neq 0$$

Thus, the system has no solution! This means that nonconstant polynomials  $f(x), g(x) \in \mathbb{F}_2[x]$  such that

$$x^3 + x + 1 = f(x)g(x)$$

do not exist! Therefore,  $x^3 + x + 1$  is irreducible. The statement now follows from the discussion from the beginning.



### Second case.

Notice that

$$(x^2 + x + 2)(x + 2) = x^3 + x + 1,$$

with  $x^2 + x + 2, x + 2 \in \mathbb{F}_3[x]$ . Thus,  $x^3 + x + 1$  is not irreducible in  $\mathbb{F}_3[x]$ , so the given quotient ring is not a field.

(NOTE: Notice that 1 is a root of  $x^3 + x + 1$ . This means that  $x - 1$  divides  $x^3 + x + 1$ . Since  $-1 = 2$  in  $\mathbb{F}_3$ , we get that  $x + 2$  divides  $x^3 + x + 1$ . This is one way how you can find divisors of polynomials above fields.)

### Result

4 of 4

(a) Show that  $x^3 + x + 1$  is irreducible in  $\mathbb{F}_2[x]$ .

(b) Notice that  $(x^2 + x + 2)(x + 2) = x^3 + x + 1$ . Conclude the rest.

### 4. a

The maximal ideals are precisely the principal ideals generated by some nonconstant irreducible polynomial.

First of all, polynomials of the first degree are clearly irreducible.

If a polynomial is of the second degree, then it is reducible if and only if it has a real root. Thus, irreducible polynomials will be such that

$$f(x) = (x - b)^2 + c^2$$

for some  $c \neq 0$ , since those polynomials are the only of the second degree which have no real roots.

If a polynomial is of higher degree, then it is reducible. Truly, it has a complex root. If that root is also real, then  $x - a$ , where  $a$  is that root, divides  $f(x)$ . If it is complex, then  $\bar{a}$  is also a root of  $f(x)$  by complex-conjugate theorem, and  $(x - a)(x - \bar{a})$  divides  $f(x)$ . Since  $f(x)$  is of degree 3 or more, we know get that  $f(x)$  is reducible.

Now we find a bijective correspondence:

$$\begin{aligned}(x - a) &\rightsquigarrow (a, 0) \\ ((x - b)^2 + c^2) &\rightsquigarrow (b, c)\end{aligned}$$

### Result

$$\begin{aligned}(x - a) &\rightsquigarrow (a, 0) \\ (x - b)^2 + c^2 &\rightsquigarrow (b, c)\end{aligned}$$

## Section 9

### 1. a



(a)

Notice that

$$y^2 + x^3 - 17 = (y^2 - 16) + (x^3 - 1) = (y - 4)(y + 4) + (x - 1)(x^2 + x + 1)$$

Therefore,

$$y^2 + x^3 - 17 \in (y - 4, x - 1)$$

Thus,

$$(y^2 + x^3 - 17, y - 4, x - 1) = (y - 4, x - 1)$$

Now let  $R = \mathbb{C}[x, y]$ . Then

$$R/(y - 4, x - 1) \approx \mathbb{C}[x, y]/(y^2 + x^3 - 17, y - 4, x - 1) = \mathbb{C}[x, y]/(y - 4, x - 1)$$

(just consider the homomorphism  $\mathbb{C}[x, y] \rightarrow R/(y - 4, x - 1)$  defined by  $p(x, y) \mapsto p(x, y) + (y^2 + x^3 - 17) + (y - 4, x - 1)$ ).

If we define

$$\varphi : \mathbb{C}[x, y] \rightarrow \mathbb{C}, \quad \varphi(f(x, y)) = f(1, 4),$$

we easily see that it is a surjective homomorphism. Now we need to find its kernel.

Let  $f(x, y)$  be such that  $\varphi(f(x, y)) = 0$ ; that is,  $f(1, 4) = 0$ . By the division algorithm,

$$f(x, y) = q(x, y)(x - 4) + r(y),$$

because  $x - 4$  is a polynomial in variable  $x$  of the first order, so  $r(y)$  will be constant with respect to  $x$ .

Furthermore,

$$r(y) = q_1(y)(y - 1) + c,$$

so

$$f(x, y) = q(x, y)(x - 4) + q_1(y)(y - 1) + c$$

Since  $c$  is constant and  $f(1, 4) = 0$ , we conclude that  $c = 0$ . Thus,

$$f(x, y) = q(x, y)(x - 4) + q_1(y)(y - 1) \in (x - 4, y - 1)$$

Thus,

$$\ker \varphi \subseteq (x - 4, y - 1)$$

The other inclusion is trivial, since clearly  $x - 4 \in \ker \varphi$  and  $y - 4 \in \ker \varphi$ , so  $(x - 4, y - 1) \subseteq \ker \varphi$  since  $(x - 4, y - 1)$  is the smallest ideal which contains them.

Thus

$$\ker \varphi = (x - 4, y - 1)$$

Finally, by the First Isomorphism Theorem,

$$\mathbb{C}[x, y]/(x - 4, y - 1) \approx \mathbb{C},$$

so  $\mathbb{C}[x, y]/(x - 4, y - 1)$  is a field! This also means that  $R/(x - 4, y - 1)$  is a field, meaning that  $(x - 4, y - 1)$  is a maximal ideal in  $R$  by Proposition 11.8.2 (b).

**(b)**

Similarly to **(a)**, we conclude that

$$R/(x + 1, y + 4) \approx \mathbb{C}[x, y]/(y^2 + x^3 - 17, x + 1, y + 4)$$

Now notice that  $x + 1 = 0 \Leftrightarrow x = -1$  and  $y + 4 = 0 \Leftrightarrow y = -4$ , but  $y^2 + x^3 - 17$  is not zero at  $x = -1$ ,  $y = -4$ . This means that these polynomials have no common zeros, so, by Corollary 11.9.4 we conclude that  $(x + 4, y + 4, y^2 + x^3 - 17) = \mathbb{C}[x, y]$ . But this means that

$$\mathbb{C}[x, y]/(y^2 + x^3 - 17, x + 1, y + 4) = 0$$

(the zero ideal), so  $R/(x + 1, y + 4) = 0$ . Thus, we must have that  $(x + 1, y + 4) = R$ , meaning that  $(x + 1, y + 4)$  is **not** a maximal ideal in  $R$ .

As in **(a)**,

$$R/(y^2, x^3 - 17) \approx \mathbb{C}[x, y]/(y^2 + x^3 - 17, y^2, x^3 - 17)$$

However,  $y^2 + x^3 - 17 \in (y^2, x^3 - 17)$ , so

$$(y^2 + x^3 - 17, y^2, x^3 - 17) = (y^2, x^3 - 17)$$

Now notice that  $y \notin (y^2, x^3 - 17)$ . Truly, suppose that  $y \in (y^2, x^3 - 17)$ . Then we can find polynomials  $a(x, y), b(x, y) \in \mathbb{C}[x, y]$  such that

$$y = a(x, y)y^2 + b(x, y)(x^3 - 17)$$

Now we evaluate these polynomials with  $y \rightsquigarrow 0$  to get

$$b(x, 0)(x^3 - 17) = 0$$

But this means that  $b(x, 0) = 0$  (since  $x^3 - 17$  is not a zero polynomial). By Lemma 11.9.7,  $y$  divides  $b(x, y)$ , so

$$b(x, y) = yc(x, y),$$

for some polynomial  $c(x, y) \in \mathbb{C}[x, y]$ . Thus,

$$y = a(x, y)y^2 + c(x, y)(x^3 - 17)y \implies y(1 - a(x, y)y - c(x, y)(x^3 - 17)) = 0$$

From this, since  $\mathbb{C}[x, y]$  is an integral domain,

$$a(x, y)y + c(x, y)(x^3 - 17) = 1$$

But this is impossible! Just consider the mapping  $x \rightsquigarrow \sqrt[3]{17}, y \rightsquigarrow 0$ , from which we get

$$0 = 1,$$

which is absurd.

Thus, we consider the ideal  $I = (y, y^2, x^3 - 17)$ , for which we have  $(y^2, x^3 - 17) \subset I$ . Suppose that  $I = \mathbb{C}[x, y]$ . Then  $1 \in I$ , so

$$1 = a(x, y)y + b(x, y)y^2 + c(x, y)(x^3 - 17)$$

Using the same mapping as before ( $x \rightsquigarrow \sqrt[3]{17}, y \rightsquigarrow 0$ ), we conclude that this is impossible. Thus,  $1 \notin I$ , so  $I \neq \mathbb{C}[x, y]$ .

So, there exists an ideal  $I$  such that

$$(y^2, x^3 - 17) \subset I \subset \mathbb{C}[x, y],$$

so  $(y^2, x^3 - 17)$  is not maximal in  $\mathbb{C}[x, y]$ .

Now we conclude by Proposition 11.8.2 (b) that  $\mathbb{C}[x, y]/(y^2, x^3 - 17)$  is not a field. But this implies that  $R/(y^2, x^3 - 17)$  is also not a field (since they are isomorphic)! Hence, again by Proposition 11.8.2. (b),  $(y^2, x^3 - 17)$  is **not** a maximal ideal in  $R$ .

## Result

(a) Yes.

(b) No.

(c) No.

## 2. a

Every polynomial with  $n$  variables and complex coefficients can be considered a continuous function  $\mathbb{C}^n \rightarrow \mathbb{C}$ . We can also restrict its domain to  $V$ ! Denote  $C = V \rightarrow \mathbb{C}$  for better visibility. So we first define a mapping

$$\varphi : \mathbb{C}[x_1, \dots, x_n] \rightarrow C$$

given by  $\varphi(f(x_1, \dots, x_n)) = f$ . It is clearly a homomorphism. Now we want to prove that  $I \subseteq K$ , where  $K = \ker \varphi$ .

Let  $f(x_1, \dots, x_n) \in I$ . Then for every  $(y_1, \dots, y_n) \in V$  we have that  $f(y_1, \dots, y_n) = 0$  by the definition of  $V$ . But this means that

$$\varphi(f(x_1, \dots, x_n)) = 0$$

(the null function)! Therefore,  $f(x_1, \dots, x_n) \in K$ , so

$$I \subseteq K$$

By the **Theorem 11.4.2 (a)** we conclude that there exists a (unique) homomorphism  $\bar{\varphi} : \mathbb{C}[x_1, \dots, x_n]/I \rightarrow C$ .

## Result

Hint: Theorem 11.4.2 (a) may prove useful.

## 3. a

Denote this set by  $S$ . We prove two inclusions.

$$\underline{S \subseteq U \times V}$$

Let  $(x, y) \in S$ . Then  $f_i(x) = 0$ , so  $x \in U$ . Similarly,  $g_j(y) = 0$ , so  $y \in V$ . Thus,  $(x, y) \in U \times V$ .

$$\underline{U \times V \subseteq S}$$

Let  $(x, y) \in U \times V$ . Then  $x \in U$  and  $y \in V$ . This means that  $f_i(x) = 0$  and  $g_j(y) = 0$ . Thus,  $(x, y) \in S$ .

## Result

20

Prove two inclusions. This follows from the definition of a product set.

## 4. a

Let  $U = \{f_i(x) = 0\}$ ,  $V = \{g_j(x) = 0\}$ .

### Union.

We will prove that

$$U \cup V = \{f_i(x)g_j(x) = 0\}$$

Denote  $S = \{f_i(x)g_j(x) = 0\}$ .

Let  $x \in U \cup V$ . Then  $x \in U$  or  $x \in V$ . If  $x \in U$ , then  $f_i(x) = 0$ , for all  $i$ , so  $f_i(x)g_j(x) = 0$  for all  $i, j$ . If  $x \in V$ , then  $g_j(x) = 0$ , for all  $j$ , so once again  $f_i(x)g_j(x) = 0$  for all  $i, j$ . Thus,  $x \in S$ , and  $U \cup V \subseteq S$ .

Let  $x \in S$ . Suppose that  $x \notin U \cup V$ . Then  $x \notin U$  and  $x \notin V$ . So, there exist some  $i, j$  such that  $f_i(x) \neq 0$  and  $g_j(x) \neq 0$ . But this means that  $f_i(x)g_j(x) \neq 0$ , which contradicts the fact that  $x \in S$ . Thus, we must have that  $x \in U \cup V$ , so  $S \subseteq U \cup V$ .

For conclusion,  $S = U \cup V$ .

### Intersection.

We will prove that

$$U \cap V = \{f_i(x) = 0, g_j(x) = 0\}$$

Denote  $T = \{f_i(x) = 0, g_j(x) = 0\}$ .

Let  $x \in U \cap V$ . Then  $x \in U$  and  $x \in V$ . This means that  $f_i(x) = 0$  and  $g_j(x) = 0$ , for all  $i, j$ . But this means that  $x \in T$ . Thus,  $U \cap V \subseteq T$ .

Let  $x \in T$ . Then specially  $f_i(x) = 0$  for all  $i$ , so  $x \in U$ . Similarly,  $g_j(x) = 0$ , for all  $j$ , so  $x \in V$ . Thus,  $x \in U \cap V$ , and  $T \subseteq U \cap V$ .

For conclusion,  $T = U \cap V$ .

### Statements.

$U \cap V \neq 0$  means that there is no solution to the equation

$$f_1(x) = \dots = f_m(x) = g_1(x) = \dots g_n(x) = 0$$

$U \cup V = \mathbb{C}^n$  means that each point in  $\mathbb{C}^n$  is either the root of every  $f_i$ , or of every  $g_j$ .

### **Result**

$$U \cup V = \{f_i(x)g_j(x) = 0\}$$

$$U \cap V = \{f_i(x) = 0, g_j(x) = 0\}$$

$U \cap V \neq 0$  means that there is no solution to the equation

$$f_1(x) = \dots = f_m(x) = g_1(x) = \dots g_n(x) = 0$$

$U \cup V = \mathbb{C}^n$  means that each point in  $\mathbb{C}^n$  is either the root of every  $f_i$ , or of every  $g_j$ .

5. a

Let  $I = (f_1, \dots, f_r)$ . Then

$$I = \{f_1(x)g_1(x) + \dots + f_r(x)g_r(x) \mid g_i(x) \in \mathbb{C}^n[x]\}$$

We will prove that

$$V = \{x \in \mathbb{C}^n \mid h(x) = 0 \text{ for all } h(x) \in I\}$$

Denote  $W = \{x \in \mathbb{C}^n \mid h(x) = 0 \text{ for all } h(x) \in I\}$ .

Let  $x_0 \in V$ . Then  $f_i(x_0) = 0$ , for  $i = 1, \dots, r$ . Now let  $h \in I$ . Thus,

$$h(x_0) = f_1(x_0)g_1(x_0) + \dots + f_r(x_0)g_r(x_0) = 0$$

Thus,  $h(x_0) = 0$  for every  $h \in I$ , hence  $x_0 \in W$ . This also means that  $V \subseteq W$ .

Now let  $x_0 \in W$ . Then  $h(x_0) = 0$  for all  $h \in I$ . But  $f_i \in I$ , so  $f_i(x_0) = 0$ ! This means that  $x_0 \in V$ , so  $W \subseteq V$ .

Thus,  $V = W$ , as required.

### **Result**

2 of 2

Prove that

$$V = \{x \in \mathbb{C}^n \mid h(x) = 0 \text{ for all } h(x) \in I\}$$

6. a



Let the variety  $V$  be

$$\{f_1 = 0, \dots, f_r = 0\}$$

We will prove the statement by induction on  $r$ .

When  $r = 1$ , then

$$V = \{f_1 = 0\},$$

which is an algebraic curve.

For illustration, let  $r = 2$ . Then

$$V = \{f_1 = 0, f_2 = 0\} = \{f_1 = 0\} \cap \{f_2 = 0\}$$

Now let  $h(t, x)$  be the greatest common divisor of  $f_1(t, x)$  and  $f_2(t, x)$ . Then

$$f_1(t, x) = h(t, x)g_1(t, x)$$

$$f_2(t, x) = h(t, x)g_2(t, x)$$

Thus,  $(t_0, x_0) \in V$  if and only if

$$h(t_0, x_0)g_1(t_0, x_0) = 0 \quad \text{and} \quad h(t_0, x_0)g_2(t_0, x_0) = 0$$

Thus, either  $h(t_0, x_0) = 0$ , or  $g_1(t_0, x_0) = 0 = g_2(t_0, x_0)$ . Since  $g_1, g_2$  do not have a common nonconstant factor, this can, by **Theorem 11.9.10** happen in only finitely many points. Thus,

$$V = \{h = 0\} \cup P,$$

where  $P$  is the set of finitely many points. Also, notice that  $\{h = 0\}$  is an algebraic curve.

Now suppose that the statement holds for  $r = k$ , and we must prove it for  $r = k + 1$ . Let

$$V = \{f_1 = 0, \dots, f_k = 0, f_{k+1} = 0\}$$

Then

$$V = \{f_1 = 0, \dots, f_k = 0\} \cap \{f_{k+1} = 0\}$$

By the induction hypothesis,  $\{f_1 = 0, \dots, f_k = 0\} = \{g = 0\} \cup P$ , where  $P$  is a set consisting of finitely many points, and  $\{g = 0\}$  is an algebraic curve. Therefore,

$$V = (\{g = 0\} \cup P) \cap \{f_{k+1} = 0\} = (\{g = 0\} \cap \{f_{k+1} = 0\}) \cup (P \cap \{f_{k+1} = 0\})$$

By our discussion when  $r = 2$ , we know that

$$\{g = 0\} \cap \{f_{k+1} = 0\} = \{h = 0\} \cup P'$$

Furthermore,  $P \cap \{f_{k+1} = 0\}$  is also clearly a set consisting of finitely many points. Thus,

$$V = \{h = 0\} \cup Q,$$

where  $\{h = 0\}$  is an algebraic curve, while  $Q$  is a set consisting of finitely many points. This completes the proof by induction.

## Result

Let

$$V = \{f_1 = 0, \dots, f_r = 0\}$$

Prove this by induction on  $r$ .

7. a



(a)

We need to solve the given system. From the second equation we get that  $y = 1 - x$ . Plugging this into the first equation:

$$y^2 - x^3 + x^2 = 1 \implies (1 - x)^2 - x^3 + x^2 = 1 \implies 1 - 2x + x^2 - x^3 + x^2 = 1$$

This reduces to

$$x^3 - 2x^2 + 2x = 0 \implies x(x^2 - 2x + 2) = 0$$

One solution is  $x_1 = 0$ . The other we obtain from the quadratic equation

$$x^2 - 2x + 2 = 0 \implies x_{2,3} = \frac{2 \pm \sqrt{4 - 8}}{2}$$

Thus,  $x_{2,3} = 1 \pm i$ .

Now we find  $y_i = 1 - x_i$ :

$$y_1 = 1, \quad y_2 = -i, \quad y_3 = i$$

Thus, all points are

$$(0, 1), \quad (1 + i, -i), \quad (1 - i, i)$$

(b)

Subtract the second equation from the first:

$$xy - y^2 = 0 \implies (x - y)y = 0$$

So,  $x - y = 0$  or  $y = 0$ . If  $y = 0$ , then both equations become  $x^2 = 1$ . Thus,  $x = \pm 1$ . So, we have two solutions for now:  $(x_1, y_1) = (1, 0)$  and  $(x_2, y_2) = (-1, 0)$ .

Now suppose that  $x - y = 0$ ; that is,  $x = y$ . Plugging these into either starting equation yields

$$3y^2 = 1$$

Thus,  $y = \pm \frac{\sqrt{3}}{3}$ . So, we have two more solutions:

$$(x_3, y_3) = \left( \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3} \right), \quad (x_4, y_4) = \left( -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right)$$

Thus, all points are

$$(1, 0), \quad (-1, 0), \quad \left( \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3} \right), \quad \left( -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right)$$

(c)

From the second equation we get that  $x \neq 0$  and  $y \neq 0$ . Also,  $y = \frac{1}{x}$ . Plugging this into first equation yields

$$\frac{1}{x^2} = x^3 \implies x^5 = 1$$

Thus, we have five solutions for  $x$ :

$$x_i = e^{(2\pi)i/5}, \quad i = 0, 1, 2, 3, 4$$

Now,  $y = \frac{1}{x}$ , so

$$y_i = e^{-(2\pi)i/5}, \quad i = 0, 1, 2, 3, 4$$

Thus, all points are

$$\left( e^{(2\pi)i/5}, e^{-(2\pi)i/5} \right), \quad i = 0, 1, 2, 3, 4$$

(d)

From the first equation we get

$$x = -y^2$$

Plugging this into the second equation we get

$$y = 0$$

Thus, the only point is

$$(0, 0)$$

### Result

- (a)  $(0, 1), \quad (1 + i, -i), \quad (1 - i, i)$
- (b)  $(1, 0), \quad (-1, 0), \quad \left( \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3} \right), \quad \left( -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right)$
- (c)  $(e^{(2\pi)i/5}, e^{-(2\pi)i/5}), \quad i = 0, 1, 2, 3, 4$
- (d)  $(0, 0)$

8. a

An ideal is defined as the subset of a ring that generalizes certain subsets of the integers like the even numbers.

[Comment](#)

## Step 2 of 3

To determine: the ideals in the polynomial ring  $\mathbb{Q}[x, y]$  contain  $x^2 + y^2 - 5$  and  $xy - 2$

For determining this first consider the theorem which states that;

Suppose  $R$  is a commutative ring with unity and  $I$  is an ideal. Then there is a bijection between the prime ideals of  $R/I$  and the prime ideals of  $R$  which contain  $I$ . This bijection is as follows where;

$$\pi: R \rightarrow R/I$$

This is the canonical surjective

If  $P \subseteq R$  is prime contains  $I$ , then  $\pi(P) \subseteq R/I$  is also prime. Likewise if  $Q \subseteq R/I$  is prime, then  $\pi^{-1}(Q)$  is prime.

Hence, by using the result of above stated theorem, consider;

$$K[x, y] / \langle xy - 2, x^2 + y^2 - 5 \rangle$$

Now, let the first generator of the ideal will be;

$$xy = 2$$

$$x = \frac{y}{2}$$

This means that  $x$  is evaluated at  $\frac{y}{2}$ . Hence, the ring is isomorphic to;

$$K[y] / \left\langle \left(\frac{y}{2}\right)^2 + y^2 - 5 \right\rangle = K[y] / \left\langle \frac{4 + y^4 - 5y^2}{y^2} \right\rangle$$

Further;

$$\frac{4 + y^4 - 5y^2}{y^2} = 0$$

$$y^4 - 5y^2 + 4 = 0$$

Further by using the completing the squares method;

$$\begin{aligned} (y^2)^2 - 2 \times y^2 \times \frac{5}{2} + \left(\frac{5}{2}\right)^2 - \left(\frac{5}{2}\right)^2 &= 0 \\ \left(y^2 - \frac{5}{2}\right)^2 &= \left(\frac{5}{2}\right)^2 \\ y^2 - \frac{5}{2} &= \frac{5}{2} \\ y^2 &= 5 \end{aligned}$$

$$y = \pm\sqrt{5}$$

So, the prime quotient ring will be considered as;

$$(y + \sqrt{5})(y - \sqrt{5})$$

Hence, the ideals are;

$$\boxed{\langle \sqrt{5}, xy - 2, x^2 + y^2 - 5 \rangle \text{ and } \langle -\sqrt{5}, xy - 2, x^2 + y^2 - 5 \rangle}$$

9. a

First note that  $g = \frac{\partial f}{\partial x}$  and  $h = \frac{\partial f}{\partial y}$  are polynomials.

Now suppose that the curve  $C$  has infinitely many singular points. This means that there are infinitely many points  $(x, y)$  such that  $f(x, y) = g(x, y) = h(x, y) = 0$ .

From Theorem 11.9.10, we now get that  $f(x, y) = g(x, y)\tilde{g}(x, y)$ , where  $\tilde{g}$  is nonconstant. But  $f$  is irreducible, so  $g$  must be constant, meaning that  $g(x, y) = 0$  for all  $(x, y) \in \mathbb{C}^2$ .

Similarly, we get that  $h(x, y) = 0$ , for all  $(x, y) \in \mathbb{C}^2$ .

However,  $f$  is now constant (since both partial derivatives of  $f$  are zero), which is impossible since  $f$  must be irreducible, and we do not consider constant polynomials as irreducible.

Therefore, there must be only finitely many points  $(x, y) \in \mathbb{C}^2$  such that  $f(x, y) = g(x, y) = h(x, y) = 0$ , hence there are only finitely many singular points.

## Result

2 of 2

Assume there are infinitely many. Notice that partial derivatives are also polynomials. Use Theorem 11.9.10 and conclude that  $f$  must be constant. Argue that this is a contradiction.

## 10. a

Suppose that  $C \neq L$ . This means that there exists some point  $(x_0, y_0)$  which is on line  $L$ , while it is not on  $C$ .

Now assume that  $a \neq 0$  (if  $a = 0$ , then we must have  $b \neq 0$  to have a line, so we proceed accordingly). We can write

$$ax + by + c = 0 \iff x = -\frac{b}{a}y - \frac{c}{a}$$

Plugging this into  $f(x, y)$  we get

$$f(x, y) = 0 \iff f\left(-\frac{b}{a}y - \frac{c}{a}, y\right) = 0$$

Since  $(x_0, y_0) \notin C$ , we see that

$$f\left(-\frac{b}{a}y_0 - \frac{c}{a}, y_0\right) \neq 0$$

This means that  $g(y) = f\left(-\frac{b}{a}y - \frac{c}{a}, y\right)$  is a polynomial with complex coefficients and with only one variable which is not a zero polynomial. Also, it is of degree  $d$ . Therefore,  $g(y) = 0$  only for  $d$  points  $y$ . But this also means that

$$f\left(-\frac{b}{a}y - \frac{c}{a}, y\right) = 0$$

also for only  $d$  points!

Since  $(x, y) \in L \cap C$  if and only if  $f\left(-\frac{b}{a}y - \frac{c}{a}, y\right) = 0$ , we conclude that  $L \cap C$  consists of only  $d$  points.

## Result

2 of 2

Write  $x = -\frac{b}{a}y - \frac{c}{a}$  from the equation of the line, plug it into  $f$ . Conclude the rest.

11. a

(a)

Let  $r$  be some other point on  $L$ , which we further consider fixed. Now we pick constants  $c_1, c_2$  such that not both are zero, and

$$c_1 f_1(r) + c_2 f_2(r) = 0$$

We can do this the following way: if  $f_1(r) = 0$ , take  $c_1 = 1, c_2 = 0$ . If  $f_1(r) \neq 0$ , take

$$c_1 = -\frac{f_2(r)}{f_1(r)}, \quad c_2 = 1$$

Now look at

$$g = c_1 f_1 + c_2 f_2$$

Since clearly  $g(p) = g(q) = g(r) = 0$ , and  $g$  must be at most a quadratic polynomial (as a linear combination of quadratic polynomials), it follows that  $g = 0$  on  $L$ . Truly, by Exercise 9.10, we have that

$$\{g(x, y) = 0\} = L$$

(since  $d = 2$ ).

Now we want to prove that  $g$  is a product of linear polynomials. Let the equation of the line be  $y = ax + b$  (if it is of the form  $x = c$ , we switch the roles of  $x$  and  $y$  in the below proof).

Since  $g = 0$  on  $L$ , we can divide  $g$  by  $y - ax - b$  to get

$$g(x, y) = h(x, y)(y - ax - b)$$

Since  $g$  is a quadratic polynomial, we get that  $h(x, y)$  is a linear polynomial. This follows from the fact that

$$\deg(pq) = \deg p + \deg q,$$

where  $\deg$  is a *degree*, while  $p$  and  $q$  are polynomials with complex coefficients. Thus, if  $h(x, y)$  would have a degree of 2 or more,  $g(x, y)$  would have to be a polynomial of degree 3 or more, which is a contradiction.

Thus,  $g$  is truly a product of two linear polynomials.

(b)

First recall that

$$g = c_1 f_1 + c_2 f_2$$

Now write  $g = g_1 g_2$ , where  $g_1, g_2$  are linear polynomials. Thus,

$$g_1 g_2 = c_1 f_1 + c_2 f_2$$

If  $(x, y)$  is a point such that  $f_1(x, y) = f_2(x, y) = 0$ , we get that

$$g_1(x, y) g_2(x, y) = 0$$

Thus, either  $g_1(x, y) = 0$ , or  $g_2(x, y) = 0$ .

Now we define

$$S_1 = \{(x, y) \mid f_1(x, y), f_2(x, y), g_1(x, y) = 0\} \quad S_2 = \{(x, y) \mid f_1(x, y), f_2(x, y), g_2(x, y) = 0\}$$

If we define  $S = C_1 \cap C_2$ , then it becomes clear that

$$S = S_1 \cup S_2$$

Now we will prove that  $S_1$  and  $S_2$  contain at most 2 points.

Since  $g_1$  is linear, it cannot divide both  $f_1$  and  $f_2$  (by the assumption of this exercise). Without loss of generality, assume that  $g_1$  does not divide  $f_1$ . This means that for the **line** ( $g_1$  is linear!)  $L = \{g_1(x, y) = 0\}$  and the curve  $C = \{f_1(x, y) = 0\}$  we have that  $L \neq C$ . Once again, by Exercise 11.9.10 we have that  $L \cap C$  consists of at most 2 points. But

$$L \cap C = \{(x, y) \mid f_1(x, y) = 0, g_1(x, y) = 0\} \supseteq S_1,$$

so  $S_1$  also consists of at most 2 points!

Similarly,  $S_2$  consists of at most 2 points. It follows that  $S = S_1 \cup S_2$  consists of at most 4 points.

## Result

3 of 3

(a) Find  $c_1, c_2$  by taking some other point  $r$  of the line  $L$  and setting  $c_1 f_1(r) + c_2 f_2(r) = 0$ .

(b) Notice that (a) is a hint for this part. Also, use Exercise 11.9.10.

## 12. a

Linear combination is an expression constructed from a set of terms by multiplying each term by a constant and adding the result.

To prove: In three ways that the three polynomials;

$$f_1 = t^2 + x^2 - 2$$

$$f_2 = tx - 1$$

$$f_3 = t^3 + 5tx^2 + 1$$

Generate the unit ideal in  $\mathbb{C}[t, x]$

Clearly there are no common zeroes

If;

$$f_2(x, t) = 0$$

Then;

$$x, t \neq 0$$

And;

$$t = x^{-1}$$

Thus;

$$\begin{aligned} f_1(x, x^{-1}) &= x^2 - 2 + x^{-2} \\ &= x^{-2}(x^4 - 2x^2 + 1) \\ &= x^{-2}(x^2 - 1)^2 \end{aligned}$$



Hence, if;

$$f(x, t) = 0$$

Then;

$$(x, t) = \pm(1, 1)$$

But;

$$f_3(1, 1) = 7$$

$$f_3(-1, -1) = -5$$

Hence;  $f_1, f_2, f_3$  have no common zeroes and by the result which states that;

Let  $U$  and  $V$  be varieties in  $\mathbb{C}^n$ . Then the union  $U \cup V$  and the intersection  $U \cap V$  are varieties.

So from the result;

$$(f_1, f_2, f_3) = (1)$$

---

[Comment](#)

---

Step 4 of 4 ^

Now, write 1 as a linear combination of  $f_1, f_2, f_3$ . Here 1 can be expressed as;

$$\frac{1}{1225} \left( \begin{aligned} &(-6t^2x - 24tx^2 - 142t^2 - 352tx + 53t - 36x)f_1 \\ &+ (12t^2x - 24tx^2 - 36x^3 + 254t^2 - 568tx + 12x^2 + 36t + 140x - 1278)f_2 \\ &+ (6tx + 12x^2 + 142t + 68x - 53)f_3 \end{aligned} \right)$$

13. a

A principal ideal is an ideal  $\mathfrak{I}$  in a ring  $R$  that is generated by only one element of the ring through multiplication by each element of  $R$ .

---

[Comment](#)

---

Step 2 of 4 ^

Consider  $\varphi: \mathbb{C}[x, y] \rightarrow \mathbb{C}[t]$  be a homomorphism that is the identity on  $\mathbb{C}$  and sends;

$$x \rightarrow x(t)$$

$$y \rightarrow y(t)$$

And, such that  $x(t)$  and  $y(t)$  are not both constant.

To prove: The kernel of  $\varphi$  is a principal ideal

Claim:  $\ker \varphi$  is principal.

If not, then  $\ker \varphi$  contains two elements  $f, g$  that do not have a common factor.

It is enough to show that they do not have a common factor in  $\mathbb{C}(x)[y]$ .

For the proof suppose that;

$$h \in \mathbb{C}(x)[y]$$

It is a common factor, then:

$$h = a^{-1}h_0 \text{ for some } a \in \mathbb{C}[x], h_0 \in \mathbb{C}[x, y] \text{ by clearing denominator.}$$

Now, assuming without loss of generality that nothing of the form  $x - \alpha$  divides  $h_0$  then  $h_0$  divides

$$f, g \text{ in } \mathbb{C}(x)[y]$$

By proposition which states that;

Let  $h(t, x)$  and  $f(t, x)$  be non-zero elements of  $\mathbb{C}[t, x]$ .

Suppose that  $h$  is not divisible by any polynomial of the form  $t - \alpha$ . If  $h$  divides  $f$  in  $\mathbb{F}[x]$  then  $h$  divides  $f$  in  $\mathbb{C}[t, x]$

So, from the above result, this divides also in  $\mathbb{C}[x, y]$

This contradicts that  $f, g$  do not have a common factor.

Thus, there exist  $a_0, b_0 \in \mathbb{C}(x)[y]$  such that;

$$a_0 f + b_0 g = 1$$

And, clearing denominators then;

$$af + bg = q \\ \in \mathbb{C}[x]$$

For some;  $a, b \in \mathbb{C}[x]$

This implies;

$$\ker \varphi \cap \mathbb{C}[x]$$

This is non-trivial but this is a contradiction for any  $g \in \ker \varphi \cap \mathbb{C}[x]$  must satisfy;

$$g(x(t)) = 0$$

For all  $t$

Hence,  $g = 0$

Therefore, the kernel of  $\varphi$  is a principal ideal

## Miscellaneous Problem

1. a

We first observe the multiplicative identity 1:

$$(1 + 1)^2 = 1 + 1 \implies (1 + 1)(1 + 1) = 1 + 1$$

Now use the distributive law:

$$(1 + 1)(1 + 1) = (1 + 1)1 + (1 + 1)1 = 1 + 1 + 1 + 1$$

Thus,

$$1 + 1 + 1 + 1 = 1 + 1$$

and

$$1 + 1 = 0$$

Now let  $r \in R$ . Then

$$r + r = 1r + 1r \stackrel{(*)}{=} (1 + 1)r = 0r = 0,$$

where in  $(*)$  we used the distributive law. Thus for every  $r \in R$  we have that

$$2r = 0,$$

so  $R$  has characteristic 2.

## Result

The statement holds: HINT: show that  $1 + 1 = 0$ .

### 2. a

We define a relation on  $S \times S$ :

$$(a, b) \sim (c, d) \text{ if and only if } ad = bc$$

We first need to prove that this is an equivalence relation.

#### Reflexive?

Let  $(a, b) \in S \times S$ . Then clearly  $(a, b) \sim (a, b)$ , since  $ab = ba$  because  $S$  is a commutative semigroup.

#### Symmetric?

Let  $(a, b), (c, d) \in S \times S$  be such that  $(a, b) \sim (c, d)$ . Then  $ad = bc$ . But now  $cb = da$ , since  $S$  is a commutative semigroup. Thus,  $(c, d) \sim (a, b)$ .

#### Transitive?

Let  $(a, b), (c, d), (e, f) \in S \times S$  be such that  $(a, b) \sim (c, d)$  and  $(c, d) \sim (e, f)$ . We need to prove that  $(a, b) \sim (e, f)$ .

First,  $(a, b) \sim (c, d)$  means that  $ad = bc$ . Furthermore,  $(c, d) \sim (e, f)$  means that  $cf = de$ .

Now  $ad = bc$  implies  $adef = bcef$ . Now we use the associativity and commutativity of  $S$  to get

$$(de)(af) = adef = bcef$$

Similarly,

$$(cf)(be) = bcef = (de)(af)$$

Thus,

$$(cf)(be) = (de)(af)$$

Furthermore,  $cf = de$ , so we can use the cancellation law:

$$be = af,$$

or

$$af = be$$

From this,  $(a, b) \sim (e, f)$ , as required.

Now we define  $G$  as a set of equivalence relations. We will identify an equivalence class  $[(a, b)]$  with its representant  $(a, b)$ . On them, we define a law of composition  $\cdot$  with

$$(a, b) \cdot (c, d) = (ac, bd)$$

To prove that it is well-defined, let  $(a', b') \sim (a, b)$ ,  $(c', d') \sim (c, d)$ . We need to prove that  $(a'c', b'd') \sim (ac, bd)$ . For now, we have that

$$a'b = b'a \quad \text{and} \quad c'd = d'c$$

From the first equality,

$$a'bc'd = b'ac'd$$

Now we use the second equality

$$b'ac'd = (b'a)(c'd) = b'ad'c$$

Thus,

$$a'bc'd = b'ad'c \implies (a'c')(bd) = (b'd')(ac)$$

(we use the associativity and commutativity of law of composition on  $S$ ). But this precisely means that

$$(a'c', b'd') \sim (ac, bd),$$

so our law of composition on  $G$  is well-defined. Furthermore, we want to prove that  $(G, \cdot)$  is a group.

#### Closure?

Let  $(a, b), (c, d) \in G$ . Then  $(a, b) \cdot (c, d) \in G$  is trivial, since  $(a, b) \cdot (c, d) = (ac, bd)$  and  $ac, bd \in S$ .

#### Associativity.

Let  $(a, b), (c, d), (e, f) \in G$ . Then

$$\begin{aligned} ((a, b) \cdot (c, d)) \cdot (e, f) &= (ac, bd) \cdot (e, f) \\ &= ((ac)e, (bd)f) \\ &\stackrel{(1)}{=} (a(ce), b(df)) \\ &= (a, b) \cdot (ce, df) \\ &= (a, b) \cdot ((c, d) \cdot (e, f)) \end{aligned}$$

(in (1) we used the associativity of  $S$ ). Thus,  $G$  is associative.

#### Identity.

We will prove that  $(1, 1)$  is an identity. Let  $(a, b) \in G$ . Then

$$(1, 1) \cdot (a, b) = (1a, 1b) = (a, b)$$

$$(a, b) \cdot (1, 1) = (a1, b1) = (a, b)$$

because  $1$  is an identity in  $S$ . Thus,  $(1, 1)$  is an identity in  $G$ .

#### Inverse.

Let  $(a, b) \in G$ . Then  $(b, a) \in G$ , and

$$(a, b) \cdot (b, a) = (ab, ba) = (ab, ab) = (1, 1)$$

(the last equality holds since  $ab = ab$ ). The second to last equality holds because  $S$  is commutative.

Similarly we get

$$(b, a) \cdot (a, b) = (1, 1)$$

Thus, every element of  $G$  is invertible.

Therefore,  $G$  is a group. We can further prove that it is also a commutative group, but it does not ask of us to do that here.

Now we want to find an *embedding* of  $S$  into  $G$ . Consider the mapping

$$\varphi : S \rightarrow G, \quad \varphi(s) = (s, 1)$$

To prove that it is a homomorphism, let  $s, t \in S$ , and

$$\varphi(st) = (st, 1) = (s, 1) \cdot (t, 1) = \varphi(s) \cdot \varphi(t)$$

Thus, it truly is a homomorphism. Furthermore, it is injective! Assume that  $s, t \in S$  are such that  $\varphi(s) = \varphi(t)$ . Then

$$(s, 1) = (t, 1),$$

which means that  $s = t$  by definition of equivalence classes and relation. Thus,  $\varphi$  is an injective homomorphism; an *embedding*.

## Result

5 of 5

Define a relation  $\sim$  on  $S \times S$  with  $(a, b) \sim (c, d)$  if and only if  $ad = bc$ . Prove that it is an equivalence relation. What can you do with its equivalence classes?

### 3. a

We first prove that  $R$  is a ring by checking all properties from the definition.

$(R, +)$  is an abelian group

Closure.

The fact that

$$(a_1, a_2, \dots) + (b_1, b_2, \dots) = (a_1 + b_1, a_2 + b_2, \dots) \in R$$

is clear.

Associativity.

Let

$$(a_1, a_2, \dots), (b_1, b_2, \dots), (c_1, c_2, \dots) \in R$$

Then

$$\begin{aligned} (a_1, a_2, \dots) + ((b_1, b_2, \dots) + (c_1, c_2, \dots)) &= (a_1, a_2, \dots) + (b_1 + c_1, b_2 + c_2, \dots) \\ &= (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2), \dots) \\ &\stackrel{(1)}{=} ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2, \dots) \\ &= (a_1 + b_1, a_2 + b_2, \dots) + (c_1, c_2, \dots) \\ &= ((a_1, a_2, \dots) + (b_1, b_2, \dots)) + (c_1, c_2, \dots) \end{aligned}$$

(In (1) we used the associativity of addition on  $\mathbb{R}$ .) Thus,  $(R, +)$  is associative.

### Identity.

Notice that

$$(a_1, a_2, \dots) + (0, 0, \dots) = (a_1, a_2, \dots)$$

and

$$(0, 0, \dots) + (a_1, a_2, \dots) = (a_1, a_2, \dots)$$

for every  $(a_1, a_2, \dots) \in R$ . Thus,  $(0, 0, \dots)$  is the additive identity.

### Inverse.

Let  $(a_1, a_2, \dots) \in R$ . Then  $(-a_1, -a_2, \dots) \in R$ , and

$$(a_1, a_2, \dots) + (-a_1, -a_2, \dots) = (0, 0, \dots)$$

$$(-a_1, -a_2, \dots) + (a_1, a_2, \dots) = (0, 0, \dots)$$

Thus, every element of  $R$  has the additive inverse.

### Commutativity.

Let  $(a_1, a_2, \dots), (b_1, b_2, \dots) \in R$ . Then

$$\begin{aligned} (a_1, a_2, \dots) + (b_1, b_2, \dots) &= (a_1 + b_1, a_2 + b_2, \dots) \\ &\stackrel{(2)}{=} (b_1 + a_1, b_2 + a_2, \dots) \\ &= (b_1, b_2, \dots) + (a_1, a_2, \dots) \end{aligned}$$

(In (2) we used the commutativity of addition in  $\mathbb{R}$ .) Thus,  $(R, +)$  is commutative.

### Conclusion.

Now we can conclude that  $(R, +)$  is an abelian group.

**$(R, \cdot)$  is closed, associative and commutative, and has an identity**

### Closure.

The fact that

$$(a_1, a_2, \dots) \cdot (b_1, b_2, \dots) = (a_1 b_1, a_2 b_2, \dots) \in R$$

is clear.

### Associativity.

Let

$$(a_1, a_2, \dots), (b_1, b_2, \dots), (c_1, c_2, \dots) \in R$$

Then

$$\begin{aligned} (a_1, a_2, \dots) \cdot ((b_1, b_2, \dots) \cdot (c_1, c_2, \dots)) &= (a_1, a_2, \dots) \cdot (b_1 c_1, b_2 c_2, \dots) \\ &= (a_1 (b_1 c_1), a_2 (b_2 c_2), \dots) \\ &\stackrel{(3)}{=} ((a_1 b_1) c_1, (a_2 b_2) c_2, \dots) \\ &= (a_1 b_1, a_2 b_2, \dots) \cdot (c_1, c_2, \dots) \\ &= ((a_1, a_2, \dots) \cdot (b_1, b_2, \dots)) \cdot (c_1, c_2, \dots) \end{aligned}$$

(In (3) we used the associativity of multiplication on  $\mathbb{R}$ .) Thus,  $(R, \cdot)$  is associative.



### Identity.

Notice that

$$(a_1, a_2, \dots) \cdot (1, 1, \dots) = (a_1, a_2, \dots)$$

and

$$(1, 1, \dots) \cdot (a_1, a_2, \dots) = (a_1, a_2, \dots)$$

for every  $(a_1, a_2, \dots) \in R$ . Thus,  $(1, 1, \dots)$  is the multiplicative identity.

### Commutativity.

Let  $(a_1, a_2, \dots), (b_1, b_2, \dots) \in R$ . Then

$$\begin{aligned}(a_1, a_2, \dots) \cdot (b_1, b_2, \dots) &= (a_1 b_1, a_2 b_2, \dots) \\ &\stackrel{(4)}{=} (b_1 a_1, b_2 a_2, \dots) \\ &= (b_1, b_2, \dots) \cdot (a_1, a_2, \dots)\end{aligned}$$

(In (4) we used the commutativity of multiplication in  $\mathbb{R}$ .) Thus,  $(R, \cdot)$  is commutative.

### Conclusion.

Now we can conclude that  $(R, \cdot)$  is closed, associative, and commutative, and it has an identity.

### Distributive Law.

Let

$$(a_1, a_2, \dots), (b_1, b_2, \dots), (c_1, c_2, \dots) \in R$$

Then

$$\begin{aligned}&((a_1, a_2, \dots) + (b_1, b_2, \dots)) \cdot (c_1, c_2, \dots) \\ &= (a_1 + b_1, a_2 + b_2, \dots) \cdot (c_1, c_2, \dots) \\ &= ((a_1 + b_1)c_1, (a_2 + b_2)c_2, \dots) \\ &\stackrel{(5)}{=} (a_1 c_1 + b_1 c_1, a_2 c_2 + b_2 c_2, \dots) \\ &= (a_1 c_1, a_2 c_2, \dots) + (b_1 c_1, b_2 c_2, \dots) \\ &= (a_1, a_2, \dots) \cdot (c_1, c_2, \dots) + (b_1, b_2, \dots) \cdot (c_1, c_2, \dots)\end{aligned}$$

(In (5) we used the distributive law of field of real numbers.) Thus, the distributive law holds.

### Conclusion.

Now we conclude that  $(R, +, \cdot)$  is a ring.

### Maximal ideals.

For  $i \in \mathbb{N}$ , define  $R_i$  as

$$R_i = \{(a_1, a_2, \dots) \in R \mid a_i = 0\}$$

To check that it is an ideal, take  $(a_1, a_2, \dots), (b_1, b_2, \dots) \in R_i$  and  $(c_1, c_2, \dots) \in R$ . Then

$$(a_1, a_2, \dots) + (b_1, b_2, \dots) = (a_1 + b_1, a_2 + b_2, \dots) \in R_i$$

(since  $a_i + b_i = 0$ ), and

$$(c_1, c_2, \dots) \cdot (a_1, a_2, \dots) = (c_1 a_1, c_2 a_2, \dots) \in R_i$$

(since  $c_i a_i = 0$ ).

Now define a mapping

$$\varphi_i : R \rightarrow \mathbb{R}, \quad \varphi_i(a_1, a_2, \dots) = a_i$$

It is a homomorphism, since

$$\begin{aligned} \varphi_i((a_1, a_2, \dots) + (b_1, b_2, \dots)) &= \varphi_i(a_1 + b_1, a_2 + b_2, \dots) \\ &= a_i + b_i \\ &= \varphi_i(a_1, a_2, \dots) + \varphi_i(b_1, b_2, \dots) \end{aligned}$$

and

$$\begin{aligned} \varphi_i((a_1, a_2, \dots) \cdot (b_1, b_2, \dots)) &= \varphi_i(a_1 b_1, a_2 b_2, \dots) \\ &= a_i b_i \\ &= \varphi_i(a_1, a_2, \dots) \cdot \varphi_i(b_1, b_2, \dots) \end{aligned}$$

Also,

$$\varphi_i(1, 1, \dots) = 1$$

Notice that it is also surjective (since we can choose  $a_i$  to be any real number).

To find its kernel, we must find all  $(a_1, a_2, \dots) \in R$  such that  $\varphi_i(a_1, a_2, \dots) = 0$ . But this means that  $a_i = 0$ , so  $(a_1, a_2, \dots) \in R_i$ . Thus,

$$\ker \varphi_i \subseteq R_i$$

The other inclusion is trivial. Thus,  $\ker \varphi_i = R_i$ .

By **the First Isomorphism Theorem**,

$$R/R_i \approx \mathbb{R}$$

Thus,  $R/R_i$  is a field (since  $\mathbb{R}$  is a field), so, by **Proposition 11.8.2 (b)** we conclude that  $R_i$  is a maximal ideal in  $R$ .

Now define

$$R_\infty = \{(a_1, a_2, \dots) \in R \mid \lim_{n \rightarrow \infty} a_n = 0\}$$

(alternatively, because of the definition of  $R$ , we can define it as a set of all  $(a_1, a_2, \dots) \in R$  such that there exists  $m \in \mathbb{N}$  such that  $a_n = 0$  for all  $n \geq m$ ).

To check that it is an ideal, take  $(a_1, a_2, \dots), (b_1, b_2, \dots) \in R_\infty$  and  $(c_1, c_2, \dots) \in R$ . Then

$$(a_1, a_2, \dots) + (b_1, b_2, \dots) = (a_1 + b_1, a_2 + b_2, \dots) \in R_\infty$$

(since  $\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n = 0$ ), and

$$(c_1, c_2, \dots) \cdot (a_1, a_2, \dots) = (c_1 a_1, c_2 a_2, \dots) \in R_\infty$$

(since  $\lim_{n \rightarrow \infty} c_n a_n = \lim_{n \rightarrow \infty} c_n \lim_{n \rightarrow \infty} a_n = 0$ ).

Thus,  $R_\infty$  is an ideal. Define a mapping

$$\varphi_\infty : R_\infty \rightarrow \mathbb{R}, \quad \varphi_\infty(a_1, a_2, \dots) = \lim_{n \rightarrow \infty} a_n$$

It is a homomorphism, since

$$\begin{aligned} \varphi_\infty((a_1, a_2, \dots) + (b_1, b_2, \dots)) &= \varphi_\infty(a_1 + b_1, a_2 + b_2, \dots) \\ &= \lim_{n \rightarrow \infty} (a_n + b_n) \\ &= \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n \\ &= \varphi_\infty(a_1, a_2, \dots) + \varphi_\infty(b_1, b_2, \dots) \end{aligned}$$

and

$$\begin{aligned} \varphi_\infty((a_1, a_2, \dots) \cdot (b_1, b_2, \dots)) &= \varphi_\infty(a_1 b_1, a_2 b_2, \dots) \\ &= \lim_{n \rightarrow \infty} a_n b_n \\ &= \lim_{n \rightarrow \infty} a_n \lim_{n \rightarrow \infty} b_n \\ &= \varphi_\infty(a_1, a_2, \dots) \cdot \varphi_\infty(b_1, b_2, \dots) \end{aligned}$$

Also,

$$\varphi_\infty(1, 1, \dots) = \lim_{n \rightarrow \infty} 1 = 1$$

Notice that it is also surjective; for  $x \in \mathbb{R}$  we have

$$\varphi_\infty(x, x, \dots) = x$$

Also it is easy to see that

$$\ker \varphi_\infty = R_\infty$$

Therefore, by the *First Isomorphism Theorem*,

$$R/R_\infty \approx \mathbb{R},$$

and, by **Proposition 11.8.2 (b)** we conclude that  $R_\infty$  is a maximal ideal in  $R$ .

All that is left to prove is that there are no other maximal ideals.

Let  $I$  be some ideal in  $R$ . If  $\lim_{n \rightarrow \infty} a_n = 0$  for all  $(a_1, a_2, \dots) \in I$ , then  $I \subseteq R_\infty$ . Thus, either  $I = R_\infty$ , or it is not a maximal ideal.

Now assume that there exists some  $(a_1, a_2, \dots) \in I$  such that  $\lim_{n \rightarrow \infty} a_n = x \neq 0$ . Then

$$(a_1, a_2, \dots) = (a_1, a_2, \dots, a_m, x, x, \dots)$$

for some  $m \in \mathbb{N}$ . (Because every element of  $R$  becomes constant at some point.)

Now suppose that  $a_1 \neq 0, a_2 \neq 0, \dots, a_m \neq 0$ . Then all entries of  $(a_1, a_2, \dots)$  are nonzero, so it is a unit, because  $a_i^{-1}$  exist and

$$(a_1, a_2, \dots)(a_1^{-1}, a_2^{-1}, \dots) = (1, 1, \dots)$$

But this means that  $(1, 1, \dots) \in I$ , so  $I = R$ ! Thus,  $I$  is not a maximal ideal.

Now suppose that for every  $i \in \{1, 2, \dots, m\}$  there exists some  $b^{(i)} = (b_1^{(i)}, b_2^{(i)}, \dots) \in I$  such that  $b_i^{(i)} \neq 0$ . Then

$$(b^{(i)})^2 = ((b_1^{(i)})^2, (b_2^{(i)})^2, \dots) \in I$$

Also,

$$a^2 = (a_1^2, a_2^2, \dots) \in I$$

Now notice that

$$a^2 + (b^{(1)})^2 + \dots + (b^{(m)})^2$$

has all entries strictly positive, so it is a unit. Once again, we conclude that  $I = R$ , so it is not a maximal ideal.

Thus, there exists some  $i$  such that  $a_i = 0$  for all  $(a_1, a_2, \dots) \in I$ . But this means that  $I \subseteq R_i$ . Thus,  $I = R_i$ , or  $I$  is not maximal.

This completes the exercise, because we conclude that  $R_i, i \in \mathbb{N}$ , and  $R_\infty$  are all maximal ideals.

## Result

To prove that it is a ring, check properties from the definition.

All maximal ideals are:

$$R_i = \{(a_1, a_2, \dots) \in R \mid a_i = 0\}, \quad i \in \mathbb{N}$$

and

$$R_\infty = \{(a_1, a_2, \dots) \in R \mid \lim_{n \rightarrow \infty} a_n = 0\}$$

## 4. a

A set  $R$  is defined as the ring which shows the binary functions with respect to addition and multiplication and also satisfies the statement that the set is abelian under addition, monoid under multiplication and is distributive under multiplication with respect to addition.

a.

To classify: The rings  $R$  that contain  $\mathbb{C}$  and have dimension 2 as vector space over  $\mathbb{C}$

Let  $\{1, r\}$  be a basis for  $R$  and let the mapping;

$$\varphi: \mathbb{C}[x] \rightarrow R$$

Be defined as;

$$1 \rightarrow 1$$

$$x \rightarrow r$$

Here,  $\varphi$  is then surjection and;

$$\ker \varphi = (f); f \in \mathbb{C}[x]$$

Since,  $\mathbb{C}[x]$  is a principal ideal domain giving;

$$R \approx \mathbb{C}[x]/(f)$$

Then by the first isomorphic theorem;

$$\deg(f) > 1$$

Since otherwise  $\{1, r\}$  would be linearly independent and since;

$$r^2 = ar + b$$

$$\in R$$

For some  $a, b \in \mathbb{C}$

Now;

$$\begin{aligned} f &= x^2 - ax - b \\ &= (x - \zeta_1)(x - \zeta_2); \zeta_1, \zeta_2 \in \mathbb{C} \end{aligned}$$

If;

$$\begin{aligned} \zeta_1 &= \zeta_2 \\ &= \zeta \end{aligned}$$

Then;

$$\begin{aligned} R &\approx \mathbb{C}[x]/(f) \\ &\approx \mathbb{C}[x]/(x^2) \end{aligned}$$

By composing with the isomorphism defined by;

$$x \rightarrow x + \zeta$$

If;

$$\zeta_1 \neq \zeta_2$$

Then;

$$(x - \zeta_1) + (x - \zeta_2) = R \text{ as ideals,}$$

Hence;

$$\begin{aligned} R &\approx \mathbb{C}[x]/(f) \\ &\approx \mathbb{C}[x]/(x - \zeta_1) \times \mathbb{C}[x]/(x - \zeta_2) \\ &\approx \mathbb{C} \times \mathbb{C} \end{aligned}$$

Hence, the two possibilities are;

$$R \approx \mathbb{C}[x]/(x^2)$$

Or;

$$\mathbb{C} \times \mathbb{C}$$

**Therefore, the rings is  $\mathbb{C} \times \mathbb{C}$**

b.

To classify: The rings  $R$  that contain  $\mathbb{C}$  and have dimension 3 as vector space over  $\mathbb{C}$

Suppose there exists  $r \in R$  such that  $\{1, r, r^2\}$  is a basis for  $R$ . Then, define the map;

$$\varphi: \mathbb{C}[x] \rightarrow R$$

Such that;

$$x \rightarrow r$$

Again;

$$\ker \varphi = (f)$$

For  $\deg f < 2$

Since, if;

$$\deg f \leq 2$$

Then  $\{1, r, r^2\}$  would not be linearly independent

Also;

$$r^3 = ar^2 + br + c; a, b, c \in \mathbb{C}$$

And so;

$$f = x^3 - ax^2 - bx - c$$

If  $f$  has one triple root  $\zeta$ , then;

$$\begin{aligned} R &\approx \mathbb{C}[x] / ((x - \zeta)^3) \\ &\approx \mathbb{C}[x] / (x^3) \end{aligned}$$

If  $f$  has a double root  $\zeta$  and a simple root  $\zeta'$  then since;

$$((x - \zeta)^2) + (x - \zeta') = R$$

As ideals, that is;

$$\begin{aligned} R &\approx \mathbb{C}[x] / ((x - \zeta)^2)(x - \zeta') \\ &\approx \mathbb{C}[x] / ((x - \zeta)^2) \times \mathbb{C}[x] / (x - \zeta') \\ &\approx \mathbb{C}[x] / (x^2) \times \mathbb{C} \end{aligned}$$

By the same argument as before, finally, if  $f$  has three distinct roots then;

$$\begin{aligned} R &\approx \mathbb{C}[x] / ((x - \zeta_1)(x - \zeta_2)(x - \zeta_3)) \\ &\approx \mathbb{C}[x] / (x - \zeta_1) \times \mathbb{C}[x] / (x - \zeta_2) \times \mathbb{C}[x] / (x - \zeta_3) \\ &\approx \mathbb{C} \times \mathbb{C} \times \mathbb{C} \end{aligned}$$

Now, suppose no  $r \in R$  exists such that  $\{1, r, r^2\}$  is a basis for  $R$ . Thus, if  $\{1, r, s\}$  is a basis for  $R$  then;

$$\begin{aligned} r^2 &= a_1 r + c_1 \\ s^2 &= b_1 s + c_2 \\ rs &= a_2 r + b_2 s + c_3 \end{aligned}$$

For some  $a_i, b_i, c_i \in \mathbb{C}$

Now assume that;

$$\begin{aligned} c_1 &= c_2 \\ &= 0 \end{aligned}$$



For changing coordinates  $r \rightarrow r + \alpha$ , this gives;

$$(r + \alpha)^2 = a_1(r + \alpha) + c_1$$

This implies;

$$r^2 = (a_1 - 2\alpha)r - (\alpha^2 + a_1\alpha - c_1)$$

And, so letting  $\alpha$  be such that;

$$\begin{aligned}\alpha^2 &= a_1\alpha - c_1 \\ &= 0\end{aligned}$$

Then;

$$r^2 = a_1r$$

Similarly for  $s$ ;

$$s^2 = b_2s$$

Then;

$$\begin{aligned}(r - zs)^2 &= r^2 + 2zrs + z^2s^2 \\ &= (a_1 + 2a_3z)r + (b_2z^2 + 2b_3z)s + 2c_3z \\ &= (A_z)(r + s) + B_z\end{aligned}$$

For some  $A_z, B_z \in \mathbb{C}$

Since;

$$\{1, r + zs, (r + zs)^2\}$$

This is linearly dependent. Thus;

$$\begin{aligned}A_z &= a_1 + 2a_3z \\ &= b_2z^2 + 2b_3z\end{aligned}$$

And;

$$B_z = 2c_3z$$

Since, the first equation must hold for all  $z$ , then;

$$\begin{aligned}a_1 &= b_2 \\ &= 0\end{aligned}$$

Hence;

$$\begin{aligned}r^2 &= s^2 \\ &= 0\end{aligned}$$

And also;

$$a_3 = b_3$$

This implies;

$$\begin{aligned}(rs)^2 &= r^2s^2 \\ &= 0\end{aligned}$$

But since;

$$\begin{aligned}(rs)^2 &= (a_3r + a_3s + c_3)^2 \\ &= 2a_3^2 + 2a_3c_3r + 2a_3c_3s + c_3^2 \\ &= 2a_3(a_3^2 + c_3)r + 2a_3(a_3^2 + c_3)s + c_3(2a_3^2 + c_3)\end{aligned}$$

That is;

$$\begin{aligned}a_3(a_3^2 + c_3) &= c_3(2a_3^2 + c_3) \\ &= 0\end{aligned}$$

If one of  $a_3, c_3$  is nonzero, then the other is also. But this is impossible and so;

$$\begin{aligned}a_3 &= b_3 \\ &= c_3 \\ &= 0\end{aligned}$$

Finally, let;

$$\varphi: \mathbb{C}[x, y] \rightarrow R$$

This be defined by;

$$1 \rightarrow 1$$

$$x \rightarrow r$$

$$y \rightarrow s$$

That is,  $\varphi$  is then surjective, giving  $R \cong \mathbb{C}[x, y]/\ker \varphi$  by the first isomorphic theorem

Now,

$$I = (x^2, y^2, xy)$$

$$\subset \ker \varphi$$

And the reverse inclusion holds since for any;

$$f \in \mathbb{C}[x, y]$$

$$f \equiv \alpha r + \beta s + \gamma \pmod{I}; \alpha, \beta, \gamma \in \mathbb{C}$$

And  $f \rightarrow 0$  in the composition;

$$\begin{aligned} \mathbb{C}[x, y]/I &\rightarrow \mathbb{C}[x, y]/\ker \varphi \\ &\rightarrow R \end{aligned}$$

If and only if;

$$\alpha = \beta$$

$$= \gamma$$

$$= 0$$

Since,  $\{1, r, s\}$  is a basis for  $R$  that is if and only if  $f \in I$

This in this case;

$$R \cong \mathbb{C}[x, y]/(x^2, y^2, xy)$$

Thus there are four possibilities for  $R$ ;

$$\boxed{\mathbb{C}[x, y]/(x^2, y^2, xy), \mathbb{C}[x]/(x^3), \mathbb{C}[x]/(x^3) \times \mathbb{C} \times \mathbb{C} \times \mathbb{C}}$$

5. a

[Image.](#)

Let  $f(x, y) \in \mathbb{C}[x, y]$ ,

$$f(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} x^i y^j$$

(of course, we consider that only finitely many  $a_{ij}$  are nonzero). Now,

$$f(x, 0) = \sum_{i=0}^{\infty} a_{i0} x^i$$

$$f(0, y) = \sum_{j=0}^{\infty} a_{0j} y^j$$

$$f(t, t) = \sum_{k=0}^{\infty} \sum_{i+j=k} a_{ij} t^k$$

Thus, from the above, if the tuple  $(P(x), Q(y), R(t))$  is in the image of  $\varphi$ , then  $p_0 = q_0 = r_0$  and  $p_1 + q_1 = r_1$ .

We will prove that the other inclusion also holds.

Let  $P(x) \in \mathbb{C}[x]$ ,  $Q(y) \in \mathbb{C}[y]$ ,  $R(t) \in \mathbb{C}[t]$  be such that  $p_0 = q_0 = r_0$  and  $p_1 + q_1 = r_1$ . (So,  $P(x)$  has coefficients  $p_i$  and so on,  $Q(y)$  has  $q_i$ ,  $R(t)$  has  $r_i$ .) We must find a polynomial  $f(x, y) \in \mathbb{C}[x, y]$  such that

$$\varphi(f(x, y)) = (P(x), Q(y), R(t))$$

This means that

$$f(x, 0) = P(x)$$

$$f(0, y) = Q(y)$$

$$f(t, t) = R(t)$$

We first consider

$$f_0(x, y) = \sum_{i=0}^{\infty} p_i x^i$$

Clearly,  $f_0(x, y) = P(x)$ .

We now want to somewhat extend  $f_0$ . We must also take care not to "ruin" our equality  $f_0(x, y) = P(x)$ . So we define

$$f_1(x, y) = f_0(x, y) + \sum_{i=1}^{\infty} q_i y^i = p_0 + \sum_{i=1}^{\infty} p_i x^i + \sum_{i=1}^{\infty} q_i y^i$$

Notice that  $f_1(x, 0) = f_0(x, 0) = P(x)$ . Also, because  $p_0 = q_0$ ,  $f_1(0, y) = Q(y)$ .

Now we want to extend  $f_1$  further. So, we want to write

$$f_2(x, y) = f_1(x, y) + g(x, y),$$

for some  $g(x, y) \in \mathbb{C}[x, y]$ , to get the final equality, while also making sure that we do not ruin previous equalities.

The easiest way to do this is to write  $g(x, y) = xyh(x, y)$ , for some  $h(x, y) \in \mathbb{C}[x, y]$ , because then

$$f_2(x, y) = f_1(x, y) + xyh(x, y),$$

so

$$f_2(x, 0) = f_1(x, 0) = P(x)$$

$$f_2(0, y) = f_1(0, y) = Q(y)$$

Of course, the only problem with this line of thinking is: can we even find appropriate  $h(x, y)$ ?

Let us try with

$$h(x, y) = \sum_{i=0}^{\infty} (r_{i+2} - p_{i+2} - q_{i+2}) x^i$$

(the reason why we picked  $i + 2$  for indices will be made apparent below).

Then

$$f_2(x, y) = p_0 + \sum_{i=1}^{\infty} p_i x^i + \sum_{i=1}^{\infty} q_i y^i + \sum_{i=0}^{\infty} (r_{i+2} - p_{i+2} - q_{i+2}) x^{i+1} y$$

Thus,

$$f_2(t, t) = p_0 + \sum_{i=1}^{\infty} p_i t^i + \sum_{i=1}^{\infty} q_i t^i + \sum_{i=0}^{\infty} (r_{i+2} - p_{i+2} - q_{i+2}) t^{i+2}$$

(now we see why we picked  $i + 2$  for indices).

Now notice that we can write the above expression a bit differently:

$$f_2(t, t) = p_0 + p_1 t + q_1 t + \sum_{i=2}^{\infty} p_i t^i + \sum_{i=2}^{\infty} q_i t^i + \sum_{i=2}^{\infty} (r_i - p_i - q_i) t^i$$

Thus,

$$f_2(t, t) = p_0 + (p_1 + q_1) t + \sum_{i=2}^{\infty} r_i t^i$$

All that is left is to use that  $r_0 = p_0$  and  $r_1 = p_1 + q_1$ . Thus,

$$f_2(t, t) = R(t)$$

Finally,

$$\varphi(f_2(x, y)) = (P(x), Q(y), R(t)),$$

so we have proven the other inclusion (regarding the image).

[Kernel.](#)

Let  $f(x, y) \in \ker \varphi$ . Set

$$f(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} x^i y^j$$

(Again, recall that we consider that only finitely many  $a_{ij}$  are nonzero.)

Then

$$(0, 0, 0) = \varphi(f(x, y)) = (f(x, 0), f(0, y), f(t, t))$$

From  $f(x, 0) = 0$  we get

$$a_{00} + a_{10}x + a_{20}x^2 + \dots = 0$$

Thus,  $a_{i0} = 0$  for all  $i \in \mathbb{N}$ . This means that

$$f(x, y) = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} a_{ij} x^i y^j = y \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} a_{ij} x^i y^{j-1}$$

Now set

$$g(x, y) = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} a_{ij} x^i y^{j-1}$$

Therefore,

$$f(x, y) = yg(x, y)$$

For simplicity,

$$g(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} b_{ij} x^i y^j$$

From  $f(0, y) = 0$ , we get

$$yg(0, y) = 0$$

Notice that  $y, g(0, y) \in \mathbb{C}[y]$ . This  $\mathbb{C}[y]$  is an integral domain, and  $y \neq 0$ , we must have that  $g(0, y) = 0$ . Similarly to before, we now get that  $g(x, y) = xh(x, y)$ , for some polynomial  $h(x, y) \in \mathbb{C}[x, y]$ . Thus,

$$f(x, y) = xyh(x, y)$$

Because of

$$f(t, t) = 0$$

and

$$f(t, t) = t^2 h(t, t),$$

we get that

$$t^2 h(t, t) = 0$$

Because  $\mathbb{C}[t]$  is an integral domain, we must have that

$$h(t, t) = 0$$

Now we perform the division algorithm; divide  $h(x, y)$  by  $x - y$  to get

$$h(x, y) = (x - y)q(x, y) + r(y)$$

(notice that  $x - y$  can be considered a polynomial in variable  $x$  with coefficients from  $\mathbb{C}[y]$ , so  $r$  will be constant with regard to  $x$ ). Now,

$$0 = h(t, t) = r(t),$$

which means that  $r(y) = 0$  (it does not matter what we call the variable, so  $y = t$  in some sense). Thus,

$$h(x, y) = (x - y)q(x, y)$$

Finally,

$$f(x, y) = xy(x - y)q(x, y),$$

which means that

$$f(x, y) \in (xy(x - y))$$

Thus,

$$\ker \varphi \subseteq (xy(x - y))$$

For the other inclusion, let  $p(x, y) \in (xy(x - y))$ . Then  $p(x, y) = xy(x - y)q(x, y)$ , for some  $q(x, y) \in \mathbb{C}[x, y]$ . We immediately get that

$$p(x, 0) = 0$$

$$p(0, y) = 0$$

$$p(t, t) = 0,$$

so

$$\varphi(p(x, y)) = (p(x, 0), p(0, y), p(t, t)) = (0, 0, 0),$$

and

$$p(x, y) \in \ker \varphi$$

Finally,  $(xy(x - y)) \subseteq \ker \varphi$ , and

$$\ker \varphi = (xy(x - y))$$

## Result

7 of 7

The image consists of all tuples  $(P(x), Q(y), R(t))$ , where  $P(x) \in \mathbb{C}[x]$ ,  $Q(y) \in \mathbb{C}[y]$ , and  $R(t) \in \mathbb{C}[t]$ ,  $P(x)$  with coefficients  $p_i$ ,  $Q(y)$  with coefficients  $q_i$ , and  $R(t)$  with coefficients  $r_i$ , such that  $p_0 = q_0 = r_0$ ,  $r_1 = p_1 + q_1$ .

The kernel is equal to the principal ideal  $(xy(x - y))$ .

## 6. a

Suppose that it lies on some algebraic curve. This means that there exists some polynomial  $f(x, y) \in \mathbb{C}[x, y]$  such that  $f(x, \sin x) = 0$ , for every  $x \in \mathbb{R}$ .

Now take some  $c \in [-1, 1]$ . Since  $\sin x = c$  for infinitely many  $x \in \mathbb{R}$ , we have that the polynomials  $f(x, y)$  and  $x - c$  have infinitely many common zeros. By Theorem 11.9.10, this means that  $f(x, y) = (x - c)g(x, y)$  for some nonconstant polynomial  $g(x, y) \in \mathbb{C}[x, y]$ . But now we can repeat this process with different  $\tilde{c} \in [-1, 1]$  to conclude that  $g(x, y) = (x - \tilde{c})h(x, y)$ ! (This is because  $f(x, y)$  has infinitely many common zeros with  $x - \tilde{c}$ , and  $(x - c)$  does not have a zero at  $\tilde{c}$ ; thus,  $g(x, y)$  must have infinitely many common zeros with  $x - \tilde{c}$ .) We can repeat this process infinitely many times (for every real number between  $-1$  and  $1$ ), and we get that  $f(x, y)$  is of infinite degree! This is clearly a contradiction.

## Result

2 of 2

Assume that it lies on some algebraic curve. HINT: Theorem 11.9.10.

## 7. a



(a)

Let  $I = (f_1, f_2, \dots, f_n)$ . Then  $f_i \in I$ , so  $f_i^2 \in I$ , for  $i = 1, 2, \dots, n$ . Furthermore, now we can conclude that

$$g = f_1^2 + \dots + f_n^2 \in I$$

since  $I$  is closed under addition.

Let  $x \in [0, 1]$ . Since  $f_i$  have no common zeros, there exists some  $j \in \{1, \dots, n\}$  such that  $f_j(x) \neq 0$ . Then we have that  $f_j(x)^2 > 0$ . Finally,

$$g(x) = f_1(x)^2 + \dots + f_j(x)^2 + \dots + f_n(x)^2 \geq f_j(x)^2 > 0$$

Thus,  $g(x) \neq 0$ . From this it follows that  $g(x) > 0$  for all  $x \in [0, 1]$ .

Now we can define

$$h : [0, 1] \rightarrow \mathbb{R}, \quad h(x) = \frac{1}{g(x)}$$

(we can define it because  $g(x) \neq 0$ ). Note that  $h$  is continuous. Furthermore,

$$(hg)(x) = h(x)g(x) = 1,$$

for all  $x \in [0, 1]$ . Thus,  $g$  is a unit, which is in the ideal  $I$ , so we have  $I = R$ , as required.

(b)

For  $y \in X$ , define

$$I_y = \{f \in R \mid f(y) = 0\}$$

First of all, we check that this is an ideal. Let  $f, g \in I_y, h \in R$ . Then

$$(f + g)(y) = f(y) + g(y) = 0 \implies f + g \in I_y$$

$$(hf)(y) = h(y)f(y) = 0 \implies hf \in I_y$$

Thus,  $I_y$  is an ideal in  $R$ .

Now we define a mapping

$$\varphi_y : R \rightarrow \mathbb{R}, \quad \varphi_y(f) = f(y)$$

We check that this is a homomorphism:

$$\varphi_y(f + g) = (f + g)(y) = f(y) + g(y) = \varphi_y(f) + \varphi_y(g)$$

$$\varphi_y(fg) = (fg)(y) = f(y)g(y) = \varphi_y(f)\varphi_y(g)$$

$$\varphi_y(\mathbf{1}) = \mathbf{1}(y) = 1$$

(we denote by  $\mathbf{1}$  the function  $\mathbf{1} : X \rightarrow \mathbb{R}, \mathbf{1}(x) = 1$ ).

Thus,  $\varphi_y$  is a homomorphism. Furthermore, it is surjective! Let  $c \in \mathbb{R}$ . Define a function  $f : X \rightarrow \mathbb{R}, f_c(x) = c$ . Then  $f_c$  is continuous, and

$$\varphi_y(f_c) = f_c(y) = c$$

We will find the kernel of  $\varphi_y$ ;  $f$  is in the kernel if and only if  $\varphi_y(f) = 0$ , which is if and only if  $f(y) = 0$ , which holds if and only if  $f \in I_y$ . Thus,  $\ker \varphi = I_y$ .

By **the First Isomorphism Theorem**, we now conclude that

$$R/I_y \approx \mathbb{R}$$

This means that  $R/I_y$  is a field, so by **Proposition 11.8.2 (b)** we conclude that  $I_y$  is a maximal ideal in  $R$ .

Now we want to prove that these are all maximal ideal in  $R$ . Let  $J$  be some ideal in  $R$ . Suppose that it is not contained in any  $I_y$ . This means that, for every  $y \in X$ , there exists some function  $f_y \in J$  such that  $f_y(y) \neq 0$ .

Now we use the fact that  $f_y$  are continuous to conclude that for every  $y \in X$ , there exists some open set  $S_y$  around  $y$  such that  $f_y(s) \neq 0$  for all  $s \in S_y$ . Clearly

$$\bigcup_{y \in X} S_y = X$$

Since  $X$  is a compact set, and  $S_y$  are open, we conclude that there exist some  $S_{y_1}, \dots, S_{y_m}$  such that

$$\bigcup_{i=1}^m S_{y_i} = X$$

Now we conclude that  $f_{y_1}, \dots, f_{y_m}$  have no common zeros in  $X$ . Truly, let  $x \in X$ . Then  $x \in S_{y_j}$  for some  $j \in \{1, \dots, m\}$ , and  $f_{y_j}(x) \neq 0$ .

Finally, we now use **(a)** to conclude that

$$(f_{y_1}, \dots, f_{y_m}) = R$$

Furthermore,

$$J \supseteq (f_{y_1}, \dots, f_{y_m}) = R$$

Thus,

$$J = R$$

which means that  $J$  is not a maximal ideal.

Finally, this means that for every ideal  $J$  we have that either  $J \subseteq I_y$  for some  $y \in X$ , or  $J = R$ , so  $I_y$  are the only maximal ideals. Now we define a bijection by

$$\Phi : \{I_y \mid y \in X\} \rightarrow X, \quad \Phi(I_y) = y$$

## Result

4 of 4

**(a)** Use the hint provided in the exercise. Show that  $g$  is a unit in  $R$ .

**(b)** Prove that  $I_y = \{f \in R \mid f(y) = 0\}$  are all maximal ideals in  $R$ .

# 12

## Chapter 12

### Section 1

1. a

Suppose that  $n$  is not a square of an integer. Now we will prove that  $n$  is not a square of any rational number.

Suppose that there exists  $q = \frac{a}{b}$  such that  $q^2 = n$ , and suppose without loss of generality that  $a, b$  are integers such that  $\gcd(a, b) = 1$ . If  $b = \pm 1$ , then  $q$  is an integer, contradicting our assumption from the beginning. Thus,  $b \neq \pm 1$ .

Use **the prime decomposition** to write  $b = \pm p_1 \cdots p_k$ , where  $p_1, \dots, p_k$  are prime. Since  $\gcd(a, b) = 1$ ,  $p_1$  does not divide  $a$  (neither do any other  $p_j$ ). Now we observe

$$q^2 = \frac{a^2}{p_1^2 \cdots p_k^2}$$

Since  $q^2 = n$ ,  $q^2$  must be an integer. Specially, we get that  $p_1$  must divide  $a^2$  (since it is in the denominator of  $q^2$ ). But  $p_1$  is prime, so  $p_1$  must divide  $a$ . However, this is a clear contradiction.

Thus, there exists no  $q \in \mathbb{Q}$  such that  $q^2 = n$ , as required.

#### Result

2 of 2

Suppose that  $q^2 = n$ . Let  $q = a/b$ , where  $a, b$  are relatively prime integers. Then  $q^2$  must be an integer (why?). Why is this impossible?

2. a

(a)

We start with

$$\frac{7}{24} = \frac{a}{8} + \frac{b}{3}$$

Multiply the equation by 24:

$$3a + 8b = 7$$

Thus,

$$3a = 7 - 8b$$

By trying various  $b$  until we get an integer divisible by 3 on the right side, we get that we can take  $b = -1$  and  $a = 5$ . Thus,

$$\frac{7}{24} = \frac{5}{8} + \frac{-1}{3}$$

(b)

Since  $u, v$  are relatively prime, there exist integer  $k, l$  such that

$$uk + vl = 1$$

Multiply this equality by  $m$  to get

$$u(km) + v(lm) = m$$

Divide the equality by  $n = uv$ :

$$\frac{m}{n} = \frac{lm}{u} + \frac{km}{v}$$

So, we set  $lm = a, km = b$ .

## Result

3 of 3

$$(a) \frac{5}{8} + \frac{-1}{3}$$

(b) There exist integers  $k, l$  such that  $uk + vl = 1$  (why?). Multiply this equality by  $m$  and conclude the rest.

3. a

(a)

We will first solve two equations:

$$nx_1 \equiv a \text{ modulo } m$$

$$mx_2 \equiv b \text{ modulo } n$$

To prove that such  $x_1$  exists, we use the fact that there exist integers  $k, l$  such that

$$km + ln = 1$$

(this is because  $m$  and  $n$  are relatively prime). Multiplying the equality by  $a$  yields

$$m(ka) + n(la) = a$$

Now,

$$a = m(ka) + n(la) \equiv n(la) \text{ modulo } m$$

Thus,

$$n(la) \equiv a \text{ modulo } m$$

so we can take  $x_1 = la$ .

Similarly, starting from  $km + ln = 1$  and multiplying it by  $b$  we obtain

$$m(kb) + n(lb) = b$$

Again,

$$b = m(kb) + n(lb) \equiv m(kb) \text{ modulo } n$$

Thus,

$$m(kb) \equiv b \text{ modulo } n$$

so we can take  $x_2 = kb$ .

(b)

Now suppose that  $y$  also satisfies

$$y \equiv a \text{ modulo } m$$

$$y \equiv b \text{ modulo } n$$

But from this we get that

$$x \equiv y \text{ modulo } m$$

$$x \equiv y \text{ modulo } n$$

Thus, there exist some integer  $k$  and  $l$  such that

$$x - y = km$$

$$x - y = ln$$

Thus,

$$km = ln$$

This means that  $n$  divides  $km$ . Since  $m$  and  $n$  are relatively prime, there exist integers  $c, d$  such that

$$cm + dn = 1$$

Multiply this equality by  $k$ :

$$c(km) + (dk)n = k$$

Plugging in  $km = ln$ , we get

$$(cl + dk)n = k$$

Thus,  $n$  divides  $k$ . For simplicity, let  $e = cl + dk$ , so  $k = en$ .

Finally, now

$$x - y = km = e(mn)$$

This means that

$$x \equiv y \text{ modulo } mn$$

Thus, for every solution  $y$  of the system of congruence relations we have that

$$x \equiv y \text{ modulo } mn$$

## Result

4 of 4

(a) Hint: write  $x = nx_1 + mx_2$ , where  $nx_1 \equiv a \text{ modulo } m$ ,  $mx_2 \equiv b \text{ modulo } n$  (of course, first prove that such  $x_1, x_2$  exist).

(b) For every other solution  $y$  we have

$$x \equiv y \text{ modulo } mn$$

4. a

## Preliminaries

We solve (a) using the same technique used in the solution of **Exercise 12.1.3**. Also, we first prove the following:

**Lemma.** If  $m$  and  $n$  are relatively prime, then from  $mx \equiv ma$  modulo  $n$  it follows that  $x \equiv a$  modulo  $n$ .

Proof. From  $mx \equiv ma$  modulo  $n$  we get that there exists an integer  $k$  such that

$$mx - ma = kn \implies m(x - a) = kn$$

Since  $m$  and  $n$  are relatively prime, there exist integers  $c, d$  such that

$$cm + dn = 1$$

Multiply the equation by  $x - a$ :

$$cm(x - a) + dn(x - a) = x - a$$

Use the fact that  $m(x - a) = nk$  to get

$$(ck + d(x - a))n = x - a$$

Thus,  $n$  divides  $x - a$ , which means that  $x \equiv a$  modulo  $n$ .

(a)

Here  $a = 3, b = 2, m = 8, n = 5$ , so we solve

$$5x_1 \equiv 3 \text{ modulo } 8$$

$$8x_2 \equiv 2 \text{ modulo } 5$$

So, we first start with  $5x_1 \equiv 3$  modulo 8. The idea is to "increase" 3 to the multiple of 5 so we can use the **Lemma** proven in the **Preliminaries**. Notice that

$$3 \equiv -5 \text{ modulo } 8$$

Thus,

$$5x_1 \equiv -5 \text{ modulo } 8$$

Dividing by 5 we get

$$x_1 \equiv -1 \text{ modulo } 8$$

So, we can take

$$x_1 = -1$$



Similarly,

$$2 \equiv 32 \text{ modulo } 5$$

Thus,

$$8x_2 \equiv 32 \text{ modulo } 5$$

Dividing by 8 we get

$$x_2 \equiv 4 \text{ modulo } 5$$

So, we can take

$$x_2 = 4$$

Now we define

$$x_0 = nx_1 + mx_2 = (5)(-1) + (8)(4) = 27$$

So, all solutions  $x$  are given by

$$x \equiv x_0 \text{ modulo } mn;$$

that is,

$$x \equiv 27 \text{ modulo } 40$$

### (b)

Here we have three congruence relations, so we solve them one by one.

From  $x \equiv 3 \text{ modulo } 15$ , we get that there exists an integer  $k$  such that

$$x - 3 = 15k,$$

so

$$x = 3 + 15k$$

Onto the next, we have

$$x \equiv 5 \text{ modulo } 8$$

On the other hand,

$$x = 3 + 15k \equiv 3 + 7k \text{ modulo } 8$$

Thus,

$$3 + 7k \equiv 5 \text{ modulo } 8 \implies 7k \equiv 2 \text{ modulo } 8$$

Since 7 and 8 are relatively prime, we can use **Lemma**. But first, we "increase" 2 to the multiple of 7:

$$2 \equiv -14 \text{ modulo } 8$$

Thus,

$$7k \equiv -14 \text{ modulo } 8$$

Dividing by 7 we get

$$k \equiv -2 \text{ modulo } 8$$

Thus, there exists an integer  $l$  such that  $k + 2 = 8l$ ; that is,  $k = 8l - 2$ .

From this,

$$x = 3 + 15k = 3 + 15(8l - 2) = -27 + 120l$$

The third congruence relation is

$$x \equiv 2 \text{ modulo } 7$$

On the other hand,

$$x = -27 + 120l \equiv 1 + l \text{ modulo } 7$$

(because  $-27 \equiv 1 \text{ modulo } 7$  and  $120 \equiv 1 \text{ modulo } 7$ , so  $120l \equiv l \text{ modulo } 7$ ). Thus,

$$1 + l \equiv 2 \text{ modulo } 7$$

This immediately yields

$$l \equiv 1 \text{ modulo } 7$$

Thus, there exists an integer  $c$  such that  $l = 1 + 7c$ . Furthermore,

$$x = -27 + 120l = -27 + 120(1 + 7c) = 93 + 840c$$

So, all solutions are of the form  $x = 93 + 840c$ , for some integer  $c$ . The converse is also true; if  $x = 93 + 840c$ , for some integer  $c$ , then  $x$  solves the three given congruence relations (this is checked manually). Thus, all solutions of the system of congruence relations are given by

$$x \equiv 93 \text{ modulo } 840$$

(c)

We solve this as we did (b). Since  $x \equiv 13 \text{ modulo } 43$ , there exists an integer  $k$  such that  $x - 13 = 43k$ . That is, we have that

$$x = 13 + 43k$$

The other congruence relation is

$$x \equiv 7 \text{ modulo } 71$$

On the other hand,

$$x = 13 + 43k \equiv 13 + 43k \text{ modulo } 71$$

Thus,

$$13 + 43k \equiv 7 \text{ modulo } 71 \implies 43k \equiv -6 \text{ modulo } 71$$

In this case, it is difficult to "increase"  $-6$  up to the multiple of 43.

So, we solve this a bit differently. We first use **the Euclidean algorithm** on 73 and 43:

$$\begin{aligned} 73 &= 43 \cdot 1 + 30 & (q_1 = 1, r_1 = 30) \\ 43 &= 30 \cdot 1 + 13 & (q_2 = 1, r_2 = 13) \\ 30 &= 13 \cdot 2 + 4 & (q_3 = 2, r_3 = 4) \\ 13 &= 4 \cdot 3 + 1 & (q_4 = 3, r_4 = 1) \\ 4 &= 1 \cdot 4 + 0 & (q_5 = 4, r_5 = 0) \end{aligned}$$

Now we write the equations a bit differently:

$$\begin{aligned} 73 &= 43 \cdot 1 + 30 \implies 30 = 73 - 43 \\ 43 &= 30 \cdot 1 + 13 \implies 13 = 43 - 30 \\ 30 &= 13 \cdot 2 + 4 \implies 4 = 30 - 13 \cdot 2 \\ 13 &= 4 \cdot 3 + 1 \implies 1 = 13 - 4 \cdot 3 \end{aligned}$$

Now we can find integers  $k, l$  such that  $73k + 43l = 1$ !

$$\begin{aligned}1 &= 13 - 4 \cdot 3 \\&= 13 - (30 - 13 \cdot 2) \cdot 3 \\&= -30 \cdot 3 + 13 \cdot 7 \\&= -30 \cdot 3 + (43 - 30) \cdot 7 \\&= 43 \cdot 7 - 30 \cdot 10 \\&= 43 \cdot 7 - (73 - 43) \cdot 10 \\&= 73 \cdot (-10) + 43 \cdot 17\end{aligned}$$

Multiplying by  $-6$ :

$$-6 = 73 \cdot 60 + 43 \cdot (-102)$$

Finally,

$$-6 = 73 \cdot 60 + 43 \cdot (-102) \equiv 43 \cdot (-102) \text{ modulo } 73$$

To simplify it further, we can notice that

$$-102 \equiv 44 \text{ modulo } 73$$

Thus,

$$43 \cdot 44 \equiv -6 \text{ modulo } 73$$

This means that we can take  $k = 44$

Now we want to prove that  $k = 44 + 73c$ , for  $c$  integer, are all solutions of  $43k \equiv -6 \text{ modulo } 73$ .

First of all, if  $k = 44 + 73c$ , then

$$43k = 43 \cdot 44 + 73 \cdot 43c \equiv 43 \cdot 44 \equiv -6 \text{ modulo } 73$$

Thus, such  $k$  are solutions.

On the other hand, suppose that  $l$  is some solution of the congruence relation; that is,  $43l \equiv -6 \text{ modulo } 73$ . Then

$$43l \equiv 43 \cdot 44 \text{ modulo } 73$$

Since 43 and 73 are relatively prime, we can use **Lemma** to conclude that

$$l \equiv 44 \text{ modulo } 73,$$

as required.

Thus,  $k = 44 + 73c$ . Furthermore,

$$x = 13 + 43k = 13 + 43(44 + 73c) = 1905 + 3139c$$

Therefore, if  $x$  solves the system of congruence relations, then  $x = 1905 + 3139c$ , for some integer  $c$ . On the other hand, all such  $x$  are solutions (this is checked manually). Thus, all solutions are given by

$$x \equiv 1905 \text{ modulo } 3139$$

## Result

8 of 8

(a)  $x \equiv 27 \text{ modulo } 40$

(b)  $x \equiv 93 \text{ modulo } 840$

(c)  $x \equiv 1905 \text{ modulo } 3139$

## 5. a

We use **the Euler Totient Theorem**: If  $a$  and  $b$  are relatively prime, then

$$a^{\varphi(b)} \equiv 1 \text{ modulo } b,$$

where  $\varphi(b)$  is a number of positive integers which are less than or equal to  $b$  and are relatively prime to  $b$ .

So, if we set  $m = \varphi(b)$ ,  $n = \varphi(a)$ , we have that

$$a^m \equiv 1 \text{ modulo } b \quad \text{and} \quad b^n \equiv 1 \text{ modulo } a$$

Furthermore, clearly  $b^n \equiv 0 \text{ modulo } b$  and  $a^m \equiv 0 \text{ modulo } a$ , so

$$a^m + b^n \equiv 1 \text{ modulo } a \quad \text{and} \quad a^m + b^n \equiv 1 \text{ modulo } b$$

From the first congruence relation, we have that

$$a^m + b^n = 1 + ka$$

for some integer  $k$ . From the second congruence relation,

$$a^m + b^n = 1 + lb$$

for some integer  $l$ . Thus, we have the equality

$$ka = lb$$

Since  $a$  and  $b$  are relatively prime, there exist integers  $c$  and  $d$  such that

$$ac + bd = 1$$

Multiplying by  $k$  yields

$$kac + kbd = k \implies lbc + kbd = k \implies (lc + kd)b = k$$

So, if we set  $e = lc + kd$ , then  $k = eb$ . This means that

$$a^m + b^n = 1 + ka = 1 + e(ab)$$

Therefore,

$$a^m + b^n \equiv 1 \text{ modulo } ab,$$

which completes the proof.

### Result

Hint: Euler's Totient Theorem.

## Section 2

### 1. a

(a)

We first check for roots. In  $\mathbb{F}_2$ , there are only two elements: 0 and 1.

$$x = 0 \implies x^3 + x^2 + x + 1 = 1 \neq 0$$

$$x = 1 \implies x^3 + x^2 + x + 1 = 4 = 0$$

( $4 = 0$  in  $\mathbb{F}_2$ ). Thus, by **Corollary 11.2.11**, we conclude that  $x - 1 = x + 1$  divides the starting polynomial. Furthermore,

$$x^3 + x^2 + x + 1 = (x + 1)(x^2 + 1)$$

Similarly,  $x^2 + 1$  has 1 for root, and

$$(x + 1)^2 = x^2 + 2x + 1 = x^2 + 1$$

Therefore,

$$x^3 + x^2 + x + 1 = (x + 1)^3$$

Now we need to check that  $x + 1$  is irreducible. Suppose that  $f(x)$  divides  $x + 1$ , so

$$x + 1 = f(x)g(x),$$

where  $g(x) \in \mathbb{F}_2[x]$ . Since  $\deg(f(x)g(x)) = \deg f(x) + \deg g(x)$  (where we denote the degree of a polynomial with  $\deg$ ), and  $\deg(x + 1) = 1$ , we get that either  $\deg f(x) = 0$  or  $\deg g(x) = 0$ . First assume that  $\deg f(x) = 0$ . But this means that  $f(x)$  is a constant nonzero polynomial in  $\mathbb{F}_2[x]$ , so  $f(x) = 1$ . However, 1 is a unit in  $\mathbb{F}_2[x]$ !

If  $\deg g(x) = 0$ , then  $g(x) = 1$ , which means that  $f(x)$  and  $x + 1$  are associates.

Thus,  $x + 1$  is irreducible.

(b)

First write

$$x^2 - 3x - 3 = x^2 + 2x + 2$$

Now it is easy to see that 1 is a root of this polynomial, so  $x - 1 = x + 4$  divides it. Now we want to find a polynomial  $p(x) \in \mathbb{F}_5[x]$  such that

$$x^2 + 2x + 2 = (x + 4)p(x)$$

Since  $\deg((x + 4)p(x)) = \deg(x + 4) + \deg p(x) = 1 + \deg p(x)$ , and  $\deg(x^2 + 2x + 2) = 2$ , we conclude that  $\deg p(x) = 1$ . Thus,  $p(x) = ax + b$ , for some  $a, b \in \mathbb{F}_5$ , hence

$$x^2 + 2x + 2 = (x + 4)(ax + b) = ax^2 + (4a + b)x + 4b$$

So, we obtain a system of equations

$$\begin{aligned} a &= 1 \\ 4a + b &= 2 \\ 4b &= 2 \end{aligned}$$

From the first equation we get  $a = 1$ .

Plug it into the second equation:

$$4 + b = 2 \implies b = -2 = 3$$



Plugging this into the third equation we get  $12 = 2$  which holds in  $\mathbb{F}_5$ . Thus,  $a = 1, b = 3$  is truly a solution of this system.

So,

$$x^2 + 2x + 2 = (x + 4)(x + 3)$$

Now we want to prove that  $x + 4$  and  $x + 3$  are irreducible. As in **(a)**, if  $f(x)$  divides  $x + 4$ , then

$$x + 4 = f(x)g(x)$$

for some  $g(x) \in \mathbb{F}_5[x]$ . As in **(a)**, we get that either  $\deg f(x) = 0$  or  $\deg g(x) = 0$ . If  $\deg f(x) = 0$ , then  $f(x) = c$ , for some  $c \in \mathbb{F}_5, c \neq 0$ . But  $c^{-1}$  exists in  $\mathbb{F}_5$  (since  $\mathbb{F}_5$  is a field), so  $h(x) = c^{-1}$  is an inverse of  $f(x)$ ! Thus,  $f(x)$  is a unit.

If  $\deg g(x) = 0$ , then  $g(x)$  is a unit, so  $f(x)$  and  $x + 4$  are associates.

Thus,  $x + 4$  is irreducible. Similarly one proves that  $x + 3$  is irreducible.

### (c)

We try all elements of  $\mathbb{F}_7$  to find roots of this polynomial:

$$x = 0 \implies x^2 + 1 = 1 \neq 0$$

$$x = 1 \implies x^2 + 1 = 2 \neq 0$$

$$x = 2 \implies x^2 + 1 = 5 \neq 0$$

$$x = 3 \implies x^2 + 1 = 10 = 3 \neq 0$$

$$x = 4 \implies x^2 + 1 = 17 = 3 \neq 0$$

$$x = 5 \implies x^2 + 1 = 26 = 5 \neq 0$$

$$x = 6 \implies x^2 + 1 = 37 = 2 \neq 0$$

Thus, this polynomial has no roots!

Now we want to prove that it is irreducible. Suppose that  $f(x)$  divides it. Then

$$x^2 + 1 = f(x)g(x)$$

First of all,  $\deg(f(x)g(x)) = \deg f(x) + \deg g(x)$ , and  $\deg(x^2 + 1) = 2$ , so  $\deg f(x) + \deg g(x) = 2$ . This means that we have three cases depending on the degree of  $f(x)$ .

If  $f(x)$  is of degree 0, then it is constant and must be nonzero for the above equality to hold. As in **(b)** we conclude that it is a unit.

If  $f(x)$  is of degree 1, then  $f(x) = ax + b$  for some  $a, b \in \mathbb{F}_7$ . But this means that  $f(a^{-1}(-b)) = 0$  ( $a^{-1}$  and  $-b$  are in  $\mathbb{F}_7$  since it is a field), so  $x^2 + 1$  would have a root! Hence, we arrived at a contradiction.

If  $f(x)$  is of degree 2, then  $g(x)$  would have to be constant, so, as in **(b)**,  $f(x)$  and  $x^2 + 1$  would have to be associates.

This proves that  $x^2 + 1$  is irreducible in  $\mathbb{F}_7$ .

### Result

$$\textbf{(a)} (x + 1)^3$$

$$\textbf{(b)} (x + 4)(x + 3)$$

$$\textbf{(c)} x^2 + 1 \text{ is already irreducible.}$$



## 2. a

We first want to factor these polynomials.

Notice that

$$\begin{aligned} x^6 + x^4 + x^3 + x^2 + x + 1 &= x^6 + x^3 + x^2 + x^4 + x + 1 \\ &= x^2(x^4 + x + 1) + (x^4 + x + 1) \\ &= (x^2 + 1)(x^4 + x + 1) \end{aligned}$$

Now we want to prove that  $p(x) = x^4 + x + 1$  is irreducible. Using the Rational Root Theorem, if  $c$  was a rational root of this polynomial, then we would have  $c = \pm 1$ . However,  $\pm 1$  is not a root of  $p(x)$ . Suppose that  $q(x) \in \mathbb{Q}[x]$  divides  $p(x)$ . Thus, if we would have

$$p(x) = q(x)r(x),$$

for  $r(x) \in \mathbb{Q}[x]$ , and  $q(x)$  and  $r(x)$  cannot be of degree 1 (because if for example  $q(x)$  was of degree 1, it would have a rational root, so  $p(x)$  would also have a rational root).

If  $q(x)$  was a nonzero constant (it clearly cannot be zero), then it would be a unit since  $\mathbb{Q}$  is a field.

If  $r(x)$  was a nonzero constant, then  $q(x)$  and  $p(x)$  would be associates.

Thus,  $q(x)$  and  $r(x)$  are of degree 2 or more.

Since  $p(x)$  is of degree 4, we conclude that  $q(x)$  and  $r(x)$  would have to be quadratic:

$$x^4 + x + 1 = (x^2 + ax + b)(x^2 + cx + d)$$

(without loss of generality we can assume that they are monic; otherwise, the product of their leading coefficients would be 1, so we can easily obtain the above form). So,

$$x^4 + x + 1 = x^4 + (a + c)x^3 + (ac + b + d)x^2 + (ad + bc)x + bd$$

Thus, we obtain a system of equations

$$\begin{aligned} a + c &= 0 \\ ac + b + d &= 0 \\ ad + bc &= 1 \\ bd &= 1 \end{aligned}$$

So,  $a = -c$  and  $b = 1/d$ . Plug this into the second equation:

$$-c^2 + 1/d + d = 0 \implies d^2 - c^2d + 1 = 0 \quad (1)$$

Plug this into the third equation:

$$-cd + c/d = 1 \implies -cd^2 + c = d$$

Thus,  $c = \frac{d}{1-d^2}$  (if  $d = \pm 1$ , then the above equality becomes  $0 = 1$ , which is a contradiction). Now plug this into (1):

$$d^2 - \frac{d^3}{(1-d^2)^2} + 1 = 0$$

Finally, after multiplying the above equality by  $(1-d^2)^2 = d^4 - 2d^2 + 1$ ,

$$d^6 - 2d^4 + d^2 - d^3 + d^4 - 2d^2 + 1 = 0 \implies d^6 - d^4 - d^3 - d^2 + 1 = 0$$

So,  $d$  is a root of a polynomial  $f(x) = x^6 - x^4 - x^3 - x^2 + 1$ . Using the Rational Root Theorem,  $d = \pm 1$ . However, now we check that for  $d = \pm 1$  the equality  $d^6 - d^4 - d^3 - d^2 + 1 = 0$  does not hold.

Thus, such rational  $d$  does not exist and, with it, neither do polynomials of degree 2 with rational coefficients  $q(x), r(x)$  such that

$$p(x) = q(x)r(x)$$

To conclude,  $p(x)$  is irreducible.

Now we factor the second polynomial from the text of the exercise. Notice that

$$\begin{aligned} x^5 + 2x^3 + x^2 + x + 1 &= x^5 + x^3 + x^2 + x^3 + x + 1 \\ &= x^2(x^3 + x + 1) + (x^3 + x + 1) \\ &= (x^2 + 1)(x^3 + x + 1) \end{aligned}$$

Thus,  $g(x) = x^2 + 1$  is a common divisor of the given polynomials. We will prove that it is their greatest common divisor.

Suppose that  $d(x)$  is some other common divisor. Since we want to find the greatest common divisor, we can take  $d(x) = (x^2 + 1)e(x)$ , because we already know that  $x^2 + 1$  is a common divisor.

Since  $d(x)$  divides  $(x^2 + 1)(x^4 + x + 1)$ , there exists some  $q(x) \in \mathbb{Q}[x]$  such that

$$d(x)q(x) = (x^2 + 1)(x^4 + x + 1) \implies (x^2 + 1)e(x)q(x) = (x^2 + 1)(x^4 + x + 1)$$

From this,

$$(x^2 + 1)(e(x)q(x) - (x^4 + x + 1)) = 0$$

Since  $\mathbb{Q}[x]$  is an integral domain, and  $x^2 + 1 \neq 0$ , we conclude that

$$e(x)q(x) - (x^4 + x + 1) = 0 \implies e(x)q(x) = x^4 + x + 1$$

Since  $x^4 + x + 1$  is irreducible, we conclude that either  $e(x)$  is constant, or  $q(x)$  is constant.

If  $e(x)$  is constant, then  $d(x) = a(x^2 + 1)$  for some  $a \in \mathbb{Q}$ , so  $d(x)$  divides  $g(x) = x^2 + 1$ . So,  $g(x)$  is still a candidate for the greatest common divisor.

If  $q(x)$  is constant, then  $e(x) = a(x^4 + x + 1)$ , so  $d(x) = a(x^2 + 1)(x^4 + x + 1)$ . Since  $d(x)$  must also divide  $(x^2 + 1)(x^3 + x + 1)$ , then

$$d(x)r(x) = (x^2 + 1)(x^3 + x + 1)$$

Similarly to before, we get that

$$a(x^4 + x + 1)r(x) = x^3 + x + 1$$

However, this equality clearly cannot hold. First of all,  $r(x) \neq 0$ , since the right side is nonzero. But this means that the degree of the left side would be higher than the degree of the right side! Thus, the equality cannot hold.

Thus,  $x^2 + 1$  is the greatest common divisor of the two given polynomials.

## Result

4 of 4

$$x^2 + 1$$

## 3. a

We check all elements of  $\mathbb{Z}/8\mathbb{Z}$ .

$$x = 0 \implies x^2 - 2 = -2 = 6 \neq 0$$

$$x = 1 \implies x^2 - 2 = -1 = 7 \neq 0$$

$$x = 2 \implies x^2 - 2 = 2 \neq 0$$

$$x = 3 \implies x^2 - 2 = 7 \neq 0$$

$$x = 4 \implies x^2 - 2 = 14 = 6 \neq 0$$

$$x = 5 \implies x^2 - 2 = 23 = 7 \neq 0$$

$$x = 6 \implies x^2 - 2 = 34 = 2 \neq 0$$

$$x = 7 \implies x^2 - 2 = 47 = 7 \neq 0$$

Thus, this polynomial has **no roots**!

## Result

No roots.

## 4. a

Suppose that there are finitely many monic irreducible polynomials. Denote them by  $p_1(x), \dots, p_n(x)$ . Furthermore, we do not consider 1 as irreducible, so  $p_i(x)$  are nonconstant.

Now we define

$$f(x) = p_1(x) \cdots p_n(x) + 1$$

First of all, it is a monic polynomial. If we denote the degree of a polynomial by  $\deg$ ,

$$\deg f(x) = \deg(p_1(x) \cdots p_n(x) + 1) = \deg(p_1(x) \cdots p_n(x)) > \deg p_i(x),$$

for all  $i = 1, 2, \dots, n$  (we used that all  $p_i(x)$  are nonconstant to get the strict inequality).

So,  $f(x) \neq p_i(x)$ , for all  $i = 1, \dots, n$ .

However,  $p_1(x), \dots, p_n(x)$  are all irreducible monic polynomials, so  $f(x)$  must be reducible.

Using the **Theorem 12.2.17 (d)**,

$$f(x) = q_1(x)q_2(x) \cdots q_k(x),$$

where  $q_i(x)$  are irreducible monic polynomials. But this means that  $q_1(x)$  is an irreducible monic polynomial which divides  $f(x)$ . Since  $p_1(x), \dots, p_n(x)$  are all irreducible monic polynomials,  $q_1(x) = p_j(x)$  for some  $j \in \{1, \dots, n\}$ .

Finally, this means that  $p_j(x)$  divides  $f(x)$ , so there exists  $q(x) \in F[x]$  such that

$$f(x) = p_j(x)q(x) \implies p_1(x) \cdots p_n(x) + 1 = p_j(x)q(x)$$

Thus,

$$p_j(x)(p_1(x) \cdots p_{j-1}(x)p_{j+1}(x) \cdots p_n(x) - q(x)) = -1$$

If we set  $h(x) = p_1(x) \cdots p_{j-1}(x)p_{j+1}(x) \cdots p_n(x) - q(x)$ , then

$$p_j(x)h(x) = -1$$

If  $h(x) = 0$ , then  $p_j(x)h(x) = 0 \neq -1$ . So,  $h(x) \neq 0$ . Now we use that  $p_j(x)$  is nonconstant to conclude that  $p_j(x)h(x)$  is also nonconstant. But this means that  $p_j(x)h(x) \neq -1$ !

Hence, we obtained a contradiction, so no  $p_j(x)$  can divide  $f(x)$ . But this contradicts the result of

**Theorem 12.2.17 (d)**. Thus,  $p_1(x), \dots, p_n(x)$  cannot be all monic irreducible polynomials.

## Result

Use the hint provided in the exercise. **Theorem 12.2.17 (d)** could be useful.

5. a

(a)

Let  $f(x)/g(x) \in C(x)$ .

If the degree of  $f(x)$  is greater than or equal to the degree of  $g(x)$ , we first divide  $f(x)$  by  $g(x)$ :

$$f(x) = g(x)q(x) + r(x)$$

Thus,

$$\frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)}$$

Now we have that the degree of  $r(x)$  is less than or equal to the degree of  $g(x)$ .

Let  $g(x)$  be

$$g(x) = (x - a)^n h(x),$$

where  $h(a) \neq 0$ . We first want to prove that there exists a decomposition of the form

$$\frac{r(x)}{(x - a)^n h(x)} = \frac{c_1}{x - a} + \dots + \frac{c_n}{(x - a)^n} + \frac{p(x)}{h(x)}, \quad (0)$$

where  $c_1, \dots, c_n \in \mathbb{C}$ , and  $p(x) \in \mathbb{C}[x]$ , and the degree of  $p(x)$  is less than the degree of  $h(x)$ .

To simplify writing, let  $y = x - a$ . Then we need to prove that

$$\frac{r(y + a)}{y^n h(y + a)} = \frac{c_1}{y} + \dots + \frac{c_n}{y^n} + \frac{p(y + a)}{h(y + a)} \quad (1)$$

Let  $\tilde{r}(y) = r(y + a)$  and so on. Notice that this substitution of variables does not change degrees of polynomials.

Multiplying the equation (1) by  $y^n \tilde{h}(y)$ , we get

$$\tilde{r}(y) = c_n \tilde{h}(y) + c_{n-1} \tilde{h}(y)y + \dots + c_1 \tilde{h}(y)y^{n-1} + \tilde{p}(y)y^n \quad (2)$$

Substitute  $y = 0$ :

$$\tilde{r}(0) = c_n \tilde{h}(0)$$

So, we choose  $c_n = \frac{\tilde{r}(0)}{\tilde{h}(0)}$ . Note that 0 is not the root of  $\tilde{h}(y)$ , since  $a$  is not the root of  $h(x)$ !



Now write (2) differently:

$$\tilde{r}(y) - c_n \tilde{h}(y) = c_{n-1} \tilde{h}(y)y + \dots + c_1 \tilde{h}(y)y^{n-1} + \tilde{p}(y)y^n$$

Since  $\tilde{r}(0) - c_n \tilde{h}(0) = 0$ ,  $y$  divides  $\tilde{r}(y) - c_n \tilde{h}(y)$ , so we write

$$y\tilde{r}_1(y) = \tilde{r}(y) - c_n \tilde{h}(y),$$

and

$$y\tilde{r}_1(y) = c_{n-1} \tilde{h}(y)y + \dots + c_1 \tilde{h}(y)y^{n-1} + \tilde{p}(y)y^n$$

Canceling  $y$  (we can do that since  $\mathbb{C}[y]$  is an integral domain), we get

$$\tilde{r}_1(y) = c_{n-1} \tilde{h}(y) + \dots + c_1 \tilde{h}(y)y^{n-2} + \tilde{p}(y)y^{n-1}$$

Repeating the process, we can find  $c_1, \dots, c_{n-1}$ . At the end, we get

$$\tilde{r}_n(y) = \tilde{p}(y)$$

So, we pick this  $\tilde{p}(y)$ . Also note that its degree is less than the degree of  $\tilde{p}(y)$ . Furthermore, its degree  $\deg \tilde{r}(y) - n$  since we had  $n$  cancelings of  $y$ ! (We denoted the degree of a polynomial by  $\deg$ .) Thus, its degree is less than the degree of  $\tilde{h}(y)$ .

Therefore, decomposition from (0) truly exists.

Now we write

$$g(x) = (x - a_1)^{n_1} (x - a_2)^{n_2} \dots (x - a_m)^{n_m}$$

(we factorize it to the linear factors; we can do that in  $\mathbb{C}[x]$ ). Thus, using the statement proven in the first two steps,

$$\frac{r(x)}{(x - a_1)^{n_1} \dots (x - a_m)^{n_m}} = \frac{c_{11}}{x - a_1} + \dots + \frac{c_{1n_1}}{(x - a_1)^{n_1}} + \frac{p(x)}{(x - a_2)^{n_2} \dots (x - a_m)^{n_m}}$$

Proceeding inductively (and using the fact that the degree of  $p(x)$  is becoming smaller and smaller), we get the desired decomposition:

$$\frac{r(x)}{(x - a_1)^{n_1} \dots (x - a_m)^{n_m}} = \frac{c_{11}}{x - a_1} + \dots + \frac{c_{1n_1}}{(x - a_1)^{n_1}} + \dots + \frac{c_{m1}}{x - a_m} + \dots + \frac{c_{mn_m}}{(x - a_m)^{n_m}}$$

Finally, this means that

$$\frac{f(x)}{g(x)} = q(x) + \frac{c_{11}}{x - a_1} + \dots + \frac{c_{1n_1}}{(x - a_1)^{n_1}} + \dots + \frac{c_{m1}}{x - a_m} + \dots + \frac{c_{mn_m}}{(x - a_m)^{n_m}},$$

and we are done.

## (b)

From (a), the set

$$\mathbb{C}[x] \cup \{1/(x - a)^i \mid i \in \mathbb{N}, a \in \mathbb{C}\}$$

spans the entire  $\mathbb{C}(x)$ . Moreover, it is easily seen to be linearly independent, so it is also a basis for this vector space.

## Result

5 of 5

(a) Let  $f(x)/g(x) \in \mathbb{C}(x)$ . First divide  $f(x)$  by  $g(x)$  to get  $f(x)/g(x) = q(x) + p(x)/g(x)$ . Now factorize  $g(x)$  and proceed inductively.

$$(b) \mathbb{C}[x] \cup \{1/(x - a)^i \mid i \in \mathbb{N}, a \in \mathbb{C}\}$$



6. a

(a)

First of all,

$$\mathbb{Z}[\omega] = \{a_0 + a_1\omega + \dots + a_n\omega^n \mid n \in \mathbb{N}, a_i \in \mathbb{Z}\}$$

However, since  $\omega^3 = 1$ , this reduces to

$$\mathbb{Z}[\omega] = \{a_0 + a_1\omega + a_2\omega^2 \mid a_0, a_1, a_2 \in \mathbb{Z}\}$$

Furthermore,

$$\begin{aligned}\omega &= \cos(2\pi/3) + i\sin(2\pi/3) = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ \omega^2 &= e^{4\pi i/3} = \cos(4\pi/3) + i\sin(4\pi/3) = -\frac{1}{2} - \frac{\sqrt{3}}{2}i\end{aligned}$$

Thus,

$$\omega^2 = -\omega - 1$$

But this means that

$$\mathbb{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbb{Z}\}$$

Now we define the size function  $\sigma$  as

$$\sigma(a + b\omega) = (a + b\omega)(a + b\bar{\omega}) = a^2 + b^2\omega\bar{\omega} + ab(\omega + \bar{\omega})$$

Since  $\bar{\omega} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i = \omega^2$ , we conclude that

$$\omega\bar{\omega} = \omega^3 = 1$$

and

$$\omega + \bar{\omega} = \omega + \omega^2 = -1$$

Thus,

$$\sigma(a + b\omega) = a^2 - ab + b^2$$

Now we want to prove that  $\sigma$  is multiplicative and that the division with remainder is possible.

Let  $a + b\omega, c + d\omega \in \mathbb{Z}[\omega]$ . Then

$$(a + b\omega)(c + d\omega) = ac + bd\omega^2 + (ad + bc)\omega$$

Now we use that  $\omega^2 = -\omega - 1$  to get

$$(a + b\omega)(c + d\omega) = (ac - bd) + (ad + bc - bd)\omega$$

Thus,

$$\sigma((a + b\omega)(c + d\omega)) = (ac - bd)^2 - (ac - bd)(ad + bc - bd) + (ad + bc - bd)^2$$

On the other hand,

$$\sigma(a + b\omega)\sigma(c + d\omega) = (a^2 - ab + b^2)(c^2 - cd + d^2)$$

Calculating,

$$\begin{aligned} & (ac - bd)^2 - (ac - bd)(ad + bc - bd) + (ad + bc - bd)^2 \\ &= b^2d^2 - abd^2 + a^2d^2 - b^2cd + abcd - a^2cd + b^2c^2 - abc^2 + a^2c^2 \end{aligned}$$

and

$$\begin{aligned} & (a^2 - ab + b^2)(c^2 - cd + d^2) \\ &= b^2d^2 - abd^2 + a^2d^2 - b^2cd + abcd - a^2cd + b^2c^2 - abc^2 + a^2c^2 \end{aligned}$$

Therefore,

$$\sigma((a + b\omega)(c + d\omega)) = \sigma(a + b\omega)\sigma(c + d\omega)$$

Therefore,  $\sigma$  is multiplicative.

We need to prove that the division with remainder holds in the sense that for every  $A, B \in \mathbb{Z}[\omega]$ ,  $B \neq 0$ , there exist  $Q, R \in \mathbb{Z}[\omega]$  such that  $A = BQ + R$  and  $R = 0$  or  $\sigma(R) < \sigma(B)$ .

Let  $A, B \in \mathbb{Z}[\omega]$ . Additionally,  $B \neq 0$ .

Consider

$$C = \frac{A}{B}$$

Now let  $A = a_0 + a_1\omega$ ,  $B = b_0 + b_1\omega$ . Then

$$C = \frac{a_0 + a_1\omega}{b_0 + b_1\omega} \cdot \frac{b_0 + b_1\bar{\omega}}{b_0 + b_1\bar{\omega}} = \frac{(a_0 + a_1\omega)(b_0 - b_1 - b_1\omega)}{b_0^2 - b_0b_1 + b_1^2}$$

Furthermore,

$$\begin{aligned} (a_0 + a_1\omega)(b_0 - b_1 - b_1\omega) &= a_0b_0 - a_0b_1 + (a_1b_0 - a_1b_1 - a_0b_1)\omega - a_1b_1\omega^2 \\ &= (a_0b_0 - a_0b_1 + a_1b_1) + (a_1b_0 - a_1b_1 - a_0b_1 + a_1b_1)\omega \\ &= (a_0b_0 - a_0b_1 + a_1b_1) + (a_1b_0 - a_0b_1)\omega \end{aligned}$$

This means that

$$C = c_0 + c_1\omega,$$

with  $c_0, c_1 \in \mathbb{Q}$ .

Now we can find integers  $m, n \in \mathbb{Z}$  such that  $|m - c_0| \leq \frac{1}{2}$ ,  $|n - c_1| \leq \frac{1}{2}$ . Define

$$Q = m + n\omega \in \mathbb{Z}[\omega]$$

Also, define

$$R = A - BQ \in \mathbb{Z}[\omega]$$

If  $R = 0$ , we are done. Otherwise, using the obvious fact that  $\sigma$  is also multiplicative on  $\mathbb{Q}[\omega]$ ,

$$\sigma(R) = \sigma(A - BQ) = \sigma(B)\sigma(A/B - Q) = \sigma(B)\sigma(C - Q)$$

Furthermore,

$$\sigma(C - Q) = \sigma((c_0 - m) + (c_1 - n)\omega) = (c_0 - m)^2 - (c_0 - m)(c_1 - n) + (c_1 - n)^2$$

By our choice for  $m, n$  we have that

$$(c_0 - m)^2 \leq \frac{1}{4},$$

$$(c_1 - n)^2 \leq \frac{1}{4},$$

and

$$(c_0 - m)(c_1 - n) \geq -\frac{1}{4}$$

The last inequality also yields

$$-(c_0 - m)(c_1 - n) \leq \frac{1}{4}$$

Therefore,

$$\sigma(C - Q) = (c_0 - m)^2 - (c_0 - m)(c_1 - n) + (c_1 - n)^2 \leq \frac{3}{4} < 1$$

Thus,

$$\sigma(R) = \sigma(B)\sigma(C - Q) < \sigma(B),$$

and our proof is complete.

## (b)

First of all,

$$\mathbb{Z}[\sqrt{-2}] = \{a_0 + a_1\sqrt{-2} + \dots + a_n(\sqrt{-2})^n \mid n \in \mathbb{N}, a_i \in \mathbb{Z}\}$$

Since  $(\sqrt{-2})^2 = -2 \in \mathbb{Z}$ , it simplifies to

$$\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} \mid a, b \in \mathbb{Z}\}$$

Now we define a function

$$\sigma(a + b\sqrt{-2}) = a^2 + 2b^2$$

To prove that it is multiplicative, let  $a + b\sqrt{-2}, c + d\sqrt{-2} \in \mathbb{Z}[\sqrt{-2}]$ . Then

$$(a + b\sqrt{-2})(c + d\sqrt{-2}) = (ac - 2bd) + (ad + bc)\sqrt{-2}$$

Thus,

$$\sigma((a + b\sqrt{-2})(c + d\sqrt{-2})) = (ac - 2bd)^2 + 2(ad + bc)^2$$

On the other hand,

$$\sigma(a + b\sqrt{-2})\sigma(c + d\sqrt{-2}) = (a^2 + 2b^2)(c^2 + 2d^2)$$

Finally,

$$(ac - 2bd)^2 + 2(ad + bc)^2 = a^2c^2 + 2a^2d^2 + 2b^2c^2 + 4b^2d^2$$

and

$$(a^2 + 2b^2)(c^2 + 2d^2) = a^2c^2 + 2a^2d^2 + 2b^2c^2 + 4b^2d^2$$

Therefore,

$$\sigma((a + b\sqrt{-2})(c + d\sqrt{-2})) = \sigma(a + b\sqrt{-2})\sigma(c + d\sqrt{-2})$$

We need to prove that the division with remainder holds in the sense that for every  $A, B \in \mathbb{Z}[\sqrt{-2}]$ ,  $B \neq 0$ , there exist  $Q, R \in \mathbb{Z}[\sqrt{-2}]$  such that  $A = BQ + R$  and  $R = 0$  or  $\sigma(R) < \sigma(B)$ . The proof is similar to the proof from **(a)**

Let  $A, B \in \mathbb{Z}[\sqrt{-2}]$ . Additionally,  $B \neq 0$ .

Consider

$$C = \frac{A}{B}$$

Now let  $A = a_0 + a_1\sqrt{-2}$ ,  $B = b_0 + b_1\sqrt{-2}$ . Then

$$C = \frac{a_0 + a_1\sqrt{-2}}{b_0 + b_1\sqrt{-2}} \cdot \frac{b_0 - b_1\sqrt{-2}}{b_0 - b_1\sqrt{-2}} = \frac{(a_0 + a_1\sqrt{-2})(b_0 - b_1\sqrt{-2})}{b_0^2 + 2b_1^2}$$

Furthermore,

$$(a_0 + a_1\sqrt{-2})(b_0 - b_1\sqrt{-2}) = a_0b_0 + 2a_1b_1 + (a_1b_0 - a_0b_1)\sqrt{-2}$$

This means that

$$C = c_0 + c_1\sqrt{-2},$$

with  $c_0, c_1 \in \mathbb{Q}$ .

Now we can find integers  $m, n \in \mathbb{Z}$  such that  $|m - c_0| \leq \frac{1}{2}$ ,  $|n - c_1| \leq \frac{1}{2}$ . Define

$$Q = m + n\sqrt{-2} \in \mathbb{Z}[\sqrt{-2}]$$

Also, define

$$R = A - BQ \in \mathbb{Z}[\sqrt{-2}]$$

If  $R = 0$ , we are done. Otherwise, using the obvious fact that  $\sigma$  is also multiplicative on  $\mathbb{Q}[\sqrt{-2}]$ ,

$$\sigma(R) = \sigma(A - BQ) = \sigma(B)\sigma(A/B - Q) = \sigma(B)\sigma(C - Q)$$

Furthermore,

$$\sigma(C - Q) = \sigma((c_0 - m) + (c_1 - n)\omega) = (c_0 - m)^2 + 2(c_1 - n)^2$$

By our choice for  $m, n$  we have that

$$(c_0 - m)^2 \leq \frac{1}{4}$$

and

$$(c_1 - n)^2 \leq \frac{1}{4}$$

The last inequality also yields

$$2(c_1 - n)^2 \leq \frac{1}{2}$$

Therefore,

$$\sigma(C - Q) = (c_0 - m)^2 + 2(c_1 - n)^2 \leq \frac{3}{4} < 1$$

Thus,

$$\sigma(R) = \sigma(B)\sigma(C - Q) < \sigma(B),$$

and our proof is complete.

## Result

Define  $\sigma$  as follows:

$$(a) \sigma(a + b\omega) = a^2 - ab + b^2.$$

$$(b) \sigma(a + b\sqrt{-2}) = a^2 + 2b^2.$$

## 7. a

Let  $d$  be the greatest common divisor of  $a$  and  $b$  in the ring of integers. Then there exist integers  $x, y$  such that

$$d = ax + by$$

Also,  $d \in \mathbb{Z}[i]$ , so it is a common divisor of  $a$  and  $b$  in  $\mathbb{Z}[i]$ .

Now let  $g$  be some common divisor of  $a$  and  $b$  in  $\mathbb{Z}[i]$ . Then  $g$  also divides  $ax$  and  $by$ . But this means that  $g$  divides  $ax + by = d$ !

Thus,  $d$  satisfies both properties from the definition of the greatest common divisor of  $a$  and  $b$  (it divides both  $a$  and  $b$ , and every other common divisor of  $a$  and  $b$  divides  $d$ ). This completes the proof.

## Result

2 of 2

Hint:  $d = ax + by$  for some integers  $x, y$ .

## 8. a

### Description of a process

Let  $a + bi$  and  $c + di$ . Then we calculate

$$\frac{a + bi}{c + di} \cdot \frac{c - di}{c - di} = \frac{ac + bd + (bc - ad)i}{c^2 + d^2} = x + yi$$

where  $x, y \in \mathbb{Q}$ . Now we take integers  $m, n$  such that  $|x - m| \leq \frac{1}{2}$  and  $|y - n| \leq \frac{1}{2}$ . Let  $q = m + ni$ . Then

$$r = (a + bi) - (c + di)q$$

### Division

Now we divide

$$\frac{4+36i}{5+i} \cdot \frac{5-i}{5-i} = \frac{(4+36i)(5-i)}{5^2+1^2} = \frac{56+176i}{26} = \frac{56}{26} + \frac{176}{26}i$$

Since

$$\frac{56}{26} \approx 2.15$$

and

$$\frac{176}{26} \approx 6.77,$$

we take  $m = 2, n = 7$ , so

$$q = 2 + 7i$$

Next,

$$r = 4 + 36i - (5+i)(2+7i) \implies r = 1 - i$$

### **Result**

$$q = 2 + 7i, r = 1 - i$$

## 9. a

Let  $I$  be some ideal in  $F[x, x^{-1}]$ . Define

$$J = I \cap F[x]$$

We will prove that  $J$  is an ideal in  $F[x]$ .

Let  $f(x), g(x) \in J, h(x) \in F[x]$ . Then  $f(x), g(x) \in I$ , so  $f(x) + g(x) \in I$  since  $I$  is an ideal (in  $F[x, x^{-1}]$ ). Moreover,  $f(x), g(x) \in F[x]$ , so  $f(x) + g(x) \in F[x]$ . Finally,  $f(x) + g(x) \in I \cap F[x] = J$ .

Similarly,  $h(x) \in F[x, x^{-1}]$  and  $f(x) \in I$  means that  $h(x)f(x) \in I$ . Moreover,  $f(x) \in F[x]$ , so  $h(x)f(x) \in I \cap F[x] = J$ .

Thus,  $J$  is an ideal in the principal ideal domain  $F[x]$ . This means that

$$J = (p(x))$$

for some  $p(x) \in F[x]$ .

Now we will prove that

$$I = (p(x))$$

Clearly  $(p(x)) = J \subseteq I$ . We need to prove the other inclusion.

Let  $q(x) \in I$  be some Laurent polynomial. If  $q(x)$  has no terms with a negative exponent, then it is in  $F[x]$ , so it is in  $J$ , which means that  $q(x) \in (p(x))$ . Now suppose that it has some term with a negative exponent. Then

$$q(x) = a_{-m}x^{-m} + \dots + a_nx^n$$



But this means that

$$x^m q(x) = a_{-m} + \dots + a_n x^{m+n} \in F[x]$$

Moreover,  $I$  is an ideal, so  $x^m q(x) \in I$ . This means that  $x^m q(x) \in J$ , so there exists some  $c(x) \in F[x]$  such that

$$x^m q(x) = c(x)p(x)$$

In  $F[x, x^{-1}]$ , we can multiply the above equality by  $x^{-m}$  to get

$$q(x) = (x^{-m}c(x))p(x)$$

So,  $x^{-m}c(x) \in F[x, x^{-1}]$ , which means that  $q(x) \in (p(x))$ . Furthermore, this means that

$$I \subseteq (p(x))$$

Finally, we obtain

$$I = (p(x)),$$

as required.

Since  $I$  was an arbitrary taken ideal in  $F[x, x^{-1}]$ , we conclude that  $F[x, x^{-1}]$  is a principal ideal domain.

## Result

3 of 3

Hint: take any ideal  $I$  in  $F[x, x^{-1}]$ . What can you say about  $I \cap F[x]$ ?

## 10. a

We will prove that  $\mathbb{R}[[t]]$  is a principle ideal domain. From **Proposition 12.2.14 (b)** will then follow that it is a unique factorization domain.

Let  $I$  be an ideal in  $\mathbb{R}[[t]]$ . If  $I = 0$ , the result is trivial. Suppose that  $I \neq 0$ .

Suppose that all elements  $f(t)$  in  $I$  are of the form

$$f(t) = a_m t^m + a_{m+1} t^{m+1} + \dots$$

(so,  $a_0 = a_1 = \dots = a_{m-1} = 0$ , and there exists some element in  $I$  with  $a_m \neq 0$ ). Of course, we can have that  $m = 0$ .

Take one element of  $I$  for which  $a_m \neq 0$ . Then we can write

$$f(t) = t^m (a_m + a_{m+1} t + \dots)$$

Since  $a_m \neq 0$ , then  $g(t) = a_m + a_{m+1} t + \dots$  is a unit in  $\mathbb{R}[[t]]$ .

Now let  $h(t) \in I$ ,

$$h(t) = b_m t^m + b_{m+1} t^{m+1} + \dots = t^m (b_m + b_{m+1} t + \dots)$$

Let  $p(t) = b_m + b_{m+1} t + \dots$  for simplicity.

Now notice that

$$h(t) = t^m p(t) = t^m g(t) g^{-1}(t) p(t) = f(t) (g^{-1}(t) p(t))$$

This means that  $h(t) \in (f(t))$ . Moreover, since  $h(t) \in I$  was taken arbitrarily,

$$I \subseteq (f(t))$$

Since  $f(t) \in I$ , the other inclusion  $(f(t)) \subseteq I$  is trivial. Thus,

$$I = (f(t))$$

This means that  $I$  is a principal ideal, and that  $\mathbb{R}[[t]]$  is a principal ideal domain.

### Result

Prove that  $\mathbb{R}[[t]]$  is a principal ideal domain and use Proposition 12.2.14 (b).

## Section 3

1. a

(a)

Let  $f(x) \in \ker \varphi$ . Then

$$\varphi(f(x)) = f(1 + \sqrt{2}) = 0$$

So,  $1 + \sqrt{2}$  is a root of  $f(x)$ . Since we can consider  $f(x)$  to be a polynomial with rational coefficients, using the **Irrational Conjugate Root Theorem** we conclude that  $1 - \sqrt{2}$  is also a root of  $f(x)$ ; that is,  $f(1 - \sqrt{2}) = 0$ . Now we divide  $f(x)$  with  $(x - (1 + \sqrt{2}))(x - (1 - \sqrt{2})) = (x - 1)^2 - 2 = x^2 - 2x - 1$  (we can do this since  $x^2 - 2x - 1$  is monic):

$$f(x) = (x^2 - 2x - 1)g(x) + r(x),$$

where  $r(x)$  has the degree 1 or less. Furthermore,

$$r(1 + \sqrt{2}) = r(1 - \sqrt{2}) = 0$$

Therefore,  $r(x)$  has 2 roots, and it is of degree 1 or less, so we must have that  $r(x) = 0$ . Thus,

$$f(x) = (x^2 - 2x - 1)g(x),$$

and

$$f(x) \in (x^2 - 2x - 1)$$

Thus,

$$\ker \varphi \subseteq (x^2 - 2x - 1)$$

For the other inclusion, let  $p(x) \in (x^2 - 2x - 1)$ . Then  $p(x) = q(x)(x^2 - 2x - 1)$ , and clearly

$$p(1 + \sqrt{2}) = 0,$$

so

$$(x^2 - 2x - 1) \subseteq \ker \varphi$$

Finally,

$$\ker \varphi = (x^2 - 2x - 1)$$

Therefore,  $\ker \varphi$  is a principal ideal.

**(b)**

Similarly to **(a)**, we conclude that if  $f(x) \in \ker \varphi$ , then  $\frac{1}{2} + \sqrt{2}$  and  $\frac{1}{2} - \sqrt{2}$  are roots of  $f(x)$ . Since

$$\left(x - \left(\frac{1}{2} + \sqrt{2}\right)\right) \left(x - \left(\frac{1}{2} - \sqrt{2}\right)\right) = \left(x - \frac{1}{2}\right)^2 - 2 = x^2 - x + \frac{1}{4} - 2 = x^2 - x - \frac{7}{4},$$

we will divide  $f(x)$  by  $x^2 - x - \frac{7}{4}$  (in  $\mathbb{Q}[x]$ , of course):

$$f(x) = \left(x^2 - x - \frac{7}{4}\right) g(x) + r(x)$$

The same arguments as in **(a)** show that  $r(x) = 0$ , so

$$f(x) = \left(x^2 - x - \frac{7}{4}\right) g(x) = (4x^2 - 4x - 7) \tilde{g}(x)$$

(where  $\tilde{g}(x) = g(x)/4 \in \mathbb{Q}[x]$ ).

Thus,  $4x^2 - 4x - 7$  is a polynomial in  $\mathbb{Z}[x]$  which divides  $f(x)$  in  $\mathbb{Q}[x]$ . It is also a primitive polynomial. By

**Theorem 12.3.6** we conclude that  $4x^2 - 4x - 7$  also divides  $f(x)$  in  $\mathbb{Z}[x]$ ! Thus,

$$f(x) = h(x)(4x^2 - 4x - 7),$$

for some  $h(x) \in \mathbb{Z}[x]$ , so

$$f(x) \in (4x^2 - 4x - 7)$$

Thus,  $\ker \varphi \subseteq (4x^2 - 4x - 7)$ . The other inclusion is proven as in **(a)**. Thus,

$$\ker \varphi = (4x^2 - 4x - 7)$$

Therefore,  $\ker \varphi$  is a principal ideal.

## Result

$$\textbf{(a)} \ker \varphi = (x^2 - 2x - 1)$$

$$\textbf{(b)} \ker \varphi = (4x^2 - 4x - 7)$$

2. a

## NOTE

The statement makes no sense unless we assume that the integer from the statement is nonzero. Thus, I assume that the said integer is nonzero.

$p(x)$  and  $q(x)$  relatively prime in  $\mathbb{Q}[x]$

Suppose that  $p(x)$  and  $q(x)$  are relatively prime in  $\mathbb{Q}[x]$ . We need to prove that, in  $\mathbb{Z}[x]$ ,  $(p(x), q(x))$  contains an integer.

Since they are relatively prime, their greatest common divisor  $d$  is 1. Also note that  $\mathbb{Q}$  is a field. By

**Theorem 12.2.17 (a)**, there exist some  $f(x), g(x) \in \mathbb{Q}[x]$  such that

$$f(x)p(x) + g(x)q(x) = 1$$

Now we multiply the above equation by the smallest common multiple of denominators of coefficients of  $f(x)$  and  $g(x)$ , which we will denote by  $k$ :

$$(kf(x))p(x) + (kg(x))q(x) = k$$

Since  $k$  is an integer, and  $kf(x), kg(x) \in \mathbb{Z}[x]$ , we conclude that  $k$  is a nonzero integer which is in  $(p(x), q(x))$ , taken as an ideal in the ring  $\mathbb{Z}[x]$ .

In  $\mathbb{Z}[x]$ ,  $(p(x), q(x))$  contains an integer.

Suppose that  $k \in \mathbb{Z}$ ,  $k \neq 0$ , such that

$$k \in (p(x), q(x)) \subseteq \mathbb{Z}[x]$$

Then there exist polynomials  $a(x), b(x) \in \mathbb{Z}[x]$  such that

$$a(x)p(x) + b(x)q(x) = k$$

Now we divide this equation by  $k \neq 0$ :

$$\frac{a(x)}{k}p(x) + \frac{b(x)}{k}q(x) = 1 \quad (1)$$

This equation makes sense in  $\mathbb{Q}[x]$ .

1 is clearly a common divisor of  $p(x)$  and  $q(x)$ . Suppose that  $c(x) \in \mathbb{Q}[x]$  is some other common divisor of  $p(x)$  and  $q(x)$ . Then from (1) we conclude that  $c(x)$  divides 1. From the definition of a greatest common divisor, we conclude that 1 is a greatest common divisor of  $p(x)$  and  $q(x)$  in  $\mathbb{Q}[x]$ , meaning that  $p(x)$  and  $q(x)$  are relatively prime.

## Result

Hint: use Theorem 12.2.17 (a).

### 3. a

Consider the provided statement to prove a version of Gauss's Lemma.

From Gauss's Lemma, let  $\mathbb{R}$  be a UFD with fraction field then,

(i) If  $f, g \in \mathbb{R}[x]$  which are primitive, then  $fg$  is also a primitive.

(ii) If  $f, g \in F[x]$  then  $c(fg) = c(f)c(g)$

The statement (ii) is followed from statement (i).

From the first statement, for each irreducible element  $p$  in  $\mathbb{R}$ , there is a ring homomorphism that is provided as below:

$$\pi_p : \mathbb{R}[x] \rightarrow \left( \frac{\mathbb{R}}{p\mathbb{R}} \right)[x]$$

The above equation is also called as reduction  $\text{mod } p$  and this is defined as,

$$\pi_p \left( \sum_{i=0}^n r_i x^i \right) = \sum_{i=0}^n \overline{r_i} x^i$$

Where  $\overline{r}$  denotes the coset  $r + p\mathbb{R}$  and this follows from a generalization of the substitution principal for polynomial rings.

An element  $f \in \mathbb{R}[x]$  is not primitive precisely, when there is an irreducible element  $p$  that divides all coefficients of  $f$ . Therefore  $f$  is primitive if and only if  $\pi_p(f) \neq 0$  for all irreducible  $p$ .

It is assume that  $f, g$  are primitive but  $fg$  is not primitive. Then, there exists some irreducible element  $p \in \mathbb{R}$  such that,

$$\begin{aligned} 0 &= \pi_p(fg) \\ &= \pi_p(f)\pi_p(g) \end{aligned}$$

As it is already shown that  $\frac{\mathbb{R}}{p\mathbb{R}}$  is an integral domain and hence  $\left( \frac{\mathbb{R}}{p\mathbb{R}} \right)[x]$  is an integral domain. Therefore, the equation implies that  $\pi_p(f) = 0$  or  $\pi_p(g) = 0$ . Both the statements are false and therefore there is an occurrence of contradiction and this **proves statement (i)**.

4. a

Suppose that

$$xy - zw = p(x, y, z, w)q(x, y, z, w)$$

If  $p(x, y, z, w)$  is constant, then it is a unit in  $\mathbb{C}[x, y, z, w]$ . If  $q(x, y, z, w)$  is constant, it is a unit, so  $xy - zw$  and  $p(x, y, z, w)$  are associates. Thus, we assume that  $p(x, y, z, w)$  and  $q(x, y, z, w)$  are not constant.

Since the polynomial  $xy - zw$  is of degree 1 in  $x$ ,  $p(x, y, z, w)$  must be of degree 0 or 1 in  $x$ . Suppose that it is of degree 1. Then  $q(x, y, z, w)$  is of degree 0 in  $x$ . (If  $p(x, y, z, w)$  is of degree 0 in  $x$ ,  $q(x, y, z, w)$  must be of degree 1, so this case is handled the same as the previous one.)

Then

$$p(x, y, z, w) = a(y, z, w)x + b(y, z, w)$$

If  $q(x, y, z, w)$  has a term with  $z$ , say  $c(y, w)z$ , then

$$p(x, y, z, w)q(x, y, z, w) = f(y, w) + a(y, z, w)c(x, y, z, w)xz$$

(with  $f(y, w)$  we denoted the sum of all other terms obtained when multiplying  $p(x, y, z, w)$  and  $q(x, y, z, w)$ ). Notice that  $f(y, w)$  must truly be constant in  $x$  and  $z$ , since  $q(x, y, z, w)$  must be constant in  $x$  and  $p(x, y, z, w)$  must be constant in  $z$  (the argument for the latter statement is the same as when we proved that  $q(x, y, z, w)$  must be constant in  $x$ ). So,

$$f(y, w) + a(y, z, w)c(x, y, z, w)xz = xy - zw$$



But this is clearly impossible.

So,  $q(x, y, z, w)$  cannot contain terms with  $z$ . It can neither contain terms with  $w$ .

Using the same arguments we used before, but switching our attention to  $z$ , we conclude that  $z$  must also be constant in  $x$ . But this means that  $z$  must be constant altogether! This is a contradiction.

Thus, the factorization

$$xy - zw = p(x, y, z, w)q(x, y, z, w)$$

where both  $p(x, y, z, w)$  and  $q(x, y, z, w)$  are nonconstant does not exist, so  $xy - zw$  is irreducible.

## Result

2 of 2

Assume that

$$xy - zw = p(x, y, z, w)q(x, y, z, w)$$

where  $p$  and  $q$  are both nonconstant. Prove that this is impossible. Conclude that  $xy - zw$  must be irreducible.

5. a

(a)

Denote by  $S$  the set of all polynomials  $p(t) \in \mathbb{C}[t]$  such that  $\frac{dp}{dt}(0) = 0$ .

$$\text{im}\psi \subseteq S$$

Let  $p(t)$  be in the image of  $\psi$ . Then  $p(t) = f(t^2, t^3)$  for some  $f(x, y) \in \mathbb{C}[x, y]$ . Applying the Chain Rule,

$$\begin{aligned} \frac{dp}{dt} &= \frac{d}{dt}(f(t^2, t^3)) \\ &= \frac{\partial f}{\partial x}(t^2, t^3) \cdot \frac{d}{dt}(t^2) + \frac{\partial f}{\partial y}(t^2, t^3) \cdot \frac{d}{dt}(t^3) \\ &= 2t \frac{\partial f}{\partial x}(t^2, t^3) + 3t^2 \frac{\partial f}{\partial y}(t^2, t^3) \end{aligned}$$

Thus,  $\frac{dp}{dt}(0) = 0$ , as required, and  $\text{im}\psi \subseteq S$ .

$$S \subseteq \text{im}\psi$$

Now suppose that  $p(t) \in \mathbb{C}[t]$  is such that  $\frac{dp}{dt}(0) = 0$ . We want to prove that it is in the image of  $\psi$ , so we must find  $f(x, y) \in \mathbb{C}[x, y]$  such that  $\psi(f(x, y)) = p(t)$ .

Let

$$p(t) = a_0 + a_1t + a_2t^2 + \dots + a_nt^n,$$

for some  $a_0, a_1, \dots, a_n \in \mathbb{C}$ ,  $n$  nonnegative integer. Then

$$\frac{dp}{dt} = a_1 + 2a_2t + \dots + na_nt^{n-1}$$

So,

$$0 = \frac{dp}{dt}(0) = a_1$$

Thus,

$$p(t) = a_0 + a_2t^2 + a_3t^3 + \dots + a_nt^n$$



Now we proceed as follows. For all  $k \geq 0$ , we define  $c_k = a_{2k}$ . This is motivated by the fact that  $x^k \rightsquigarrow t^{2k}$  under  $\psi$ . So, we took care of terms with  $t^m$  where  $m$  is even. When  $m$  is odd, it is not 1 (because  $a_1 = 0$ ), so  $m \geq 3$ . We write it in the form  $m = 2l + 3$ , where  $l$  is a nonnegative integer, and define  $d_l = a_{2l+3}$ . This is motivated by the fact that  $x^l y \rightsquigarrow t^{2l+3}$  under  $\psi$ .

Finally, define

$$f(x, y) = \sum_{k=0}^{\infty} c_k x^k + \sum_{l=0}^{\infty} d_l x^l y$$

Since by construction of  $c_k$  and  $d_l$  only finitely many of them are nonzero, we conclude that  $f(x, y) \in \mathbb{C}[x, y]$ . Furthermore, now it is clear that

$$\psi(f(x, y)) = p(t)$$

Thus,  $S \subseteq \text{im}\psi$ .

### Kernel

We want to find a relation between  $x$  and  $y$ :

$$x = t^2 - t \quad y = t^3 - t^2$$

So,

$$y = t(t^2 - t) = tx \implies t = \frac{y}{x}$$

Now plug it into  $x = t^2 - t$ :

$$x = \frac{y^2}{x^2} - \frac{y}{x}$$

After multiplying by  $x^2$ ,

$$x^3 - y^2 + xy = 0$$

Define  $g(x, y) = x^3 - y^2 + xy$ . Now it is easy to see that  $\varphi(g(x, y)) = 0$ . Thus,

$$(x^3 - y^2 + xy) \subseteq \ker \varphi$$

Now we want to prove that the other inclusion also holds.

Let  $f(x, y) \in \ker \varphi$ , and divide it by  $x^3 - y^2 + xy$ :

$$f(x, y) = (x^3 - y^2 + xy)q(x, y) + r(x, y)$$

Since  $x^3 - y^2 + xy$  is of degree 3 in  $x$ , we conclude that

$$r(x, y) = a(y)x^2 + b(y)x + c(y),$$

where  $a(y), b(y), c(y) \in \mathbb{C}[y]$ . Furthermore,

$$0 = \varphi(f(x, y)) \stackrel{(*)}{=} r(t^2 - t) = a(t^3 - t^2)(t^2 - t)^2 + b(t^3 - t^2)(t^2 - t) + c(t^3 - t^2),$$

where  $(*)$  holds because  $\varphi((x^3 - y^2 + xy)q(x, y)) = 0$ .

Now let  $k, l, m$  be the lowest powers of  $y$  in  $a(y), b(y), c(y)$ . Suppose that  $a(y) \neq 0, b(y) \neq 0, c(y) \neq 0$ . Then the lowest power of  $t$  in  $a(t^3 - t^2)(t^2 - t)^2$  is  $2k + 2$ , the lowest power of  $t$  in  $b(t^3 - t^2)(t^2 - t)$  is  $2l + 1$ , while the lowest power of  $t$  in  $c(t^3 - t^2)$  is  $2m$ .

Since we have

$$a(t^3 - t^2)(t^2 - t)^2 + b(t^3 - t^2)(t^2 - t) + c(t^3 - t^2) = 0, \quad (1)$$

the term with the lowest power in  $c(t^3 - t^2)$  must be canceled. We will prove that the terms with the lowest powers in  $c(t^3 - t^2)$  and  $a(t^3 - t^2)(t^2 - t)^2$  must cancel each other.

First of all,  $2m$  is even, so if some term of  $b(t^3 - t^2)(t^2 - t)$  has power of  $2m$ , it cannot be the term with the lowest power (since  $2l + 1$  is odd). But this means that the term with the lowest power in  $b(t^3 - t^2)(t^2 - t)$  must be canceled by some term of  $a(t^3 - t^2)(t^2 - t)^2$ . Furthermore, it cannot be the term with the lowest power in  $a(t^3 - t^2)(t^2 - t)^2$ , since  $2l + 1$  is odd, while  $2k + 2$  is even. But this means that we cannot cancel the term with the lowest power in  $a(t^3 - t^2)(t^2 - t)^2$ ! Hence, we obtained a contradiction.

This means that  $b(t^3 - t^2)(t^2 - t)$  has no terms with the power of  $2m$ , so the term with the lowest power in  $c(t^3 - t^2)$  must be canceled by some term of  $a(t^3 - t^2)(t^2 - t)^2$ . Applying the same proof as above on  $a(t^3 - t^2)(t^2 - t)^2$ , we conclude that the terms with the lowest powers in  $a(t^3 - t^2)(t^2 - t)^2$  and  $c(t^3 - t^2)$  must cancel each other.

But this means that we must have that  $2m = 2k + 2$ , so  $m = k + 1$ . We can divide the equation (1) by  $(t^3 - t^2)^k$ , so without loss of generality we assume that  $k = 0, m = 1$ . Hence

$$a_k(t^2 - t)^2 + c_m(t^3 - t) = 0,$$

and

$$a_k t^4 - 2a_k t^3 + a_k t^2 + c_m t^3 - c_m t = 0$$

But this means that  $a_k = c_m = 0$ , which contradicts the fact that  $k$  and  $m$  are lowest powers of  $y$  in  $a(y)$  and  $c(y)$ , respectively!

Thus, we must have that  $a(y) = c(y) = 0$ . Furthermore, (1) now shows that  $b(y) = 0$ . Finally,

$$f(x, y) = q(x, y)(x^3 - y^2 + xy) \implies f(x, y) \in (x^3 - y^2 + xy),$$

thus

$$\ker \varphi \subseteq (x^3 - y^2 + xy)$$

Finally,

$$\boxed{\ker \varphi = (x^3 - y^2 + xy)}$$

### Image

Let  $p(t)$  be in the image of  $\varphi$ . Then there exists some  $f(x, y) \in \mathbb{C}[x, y]$  such that

$$\varphi(f(x, y)) = p(t) \implies f(t^2 - t, t^3 - t^2) = p(t)$$

So,

$$p(0) = f(0, 0)$$

and

$$p(1) = f(1 - 1, 1 - 1) = f(0, 0),$$

so

$$p(0) = p(1)$$

To show the other inclusion, suppose that  $p(t) \in \mathbb{C}[t]$  is such that  $p(0) = p(1)$ . Define

$$\tilde{p}(t) = p(t) - p(0)$$

Then  $\tilde{p}(1) = \tilde{p}(0) = 0$ , hence  $t(t - 1)$  divides  $\tilde{p}(t)$ :

$$\tilde{p}(t) = t(t - 1)q(t)$$

for some  $q(t) \in \mathbb{C}[t]$ . Let

$$q(t) = a_0 + a_1t + \dots + a_nt^n$$

Thus,

$$p(t) = a_0t(t - 1) + a_1t^2(t - 1) + \dots + a_nt^{n+1}(t - 1)$$

We will prove that  $t^i(t - 1)^j$  is in the image of  $\varphi$  if  $j \leq i \leq 2j$ . Since

$$x \rightsquigarrow t^2 - t = t(t - 1)$$

$$y \rightsquigarrow t^3 - t^2 = t^2(t - 1),$$

we conclude that

$$x^a y^b \rightsquigarrow t^{a+2b}(t - 1)^{a+b}$$

We will prove that we can find such  $a, b$  so we have

$$x^a y^b \rightsquigarrow t^i(t - 1)^j$$

We need to have

$$a + 2b = i$$

$$a + b = j$$

So, after subtracting the second equation from the first,

$$b = i - j$$

Since  $i \geq j$ , this makes sense (meaning that  $b$  is a nonnegative integer). Plugging this into the second equation,

$$a = j - b = 2j - i$$

Since  $i \leq 2j$ , this once again makes sense.

Thus,

$$x^{2j-i}y^{i-j} \rightsquigarrow t^i(t-1)^j$$

Now suppose that  $i > 2j$  (as we will have in  $t^k(t-1)$ ). Then we can write

$$t^i(t-1)^j = t^{i-1}(t-1+1)(t-1)^j = t^{i-1}(t-1)^{j+1} + t^{i-1}(t-1)^j$$

So, in each step of this process we get closer to the condition  $i \leq 2j$ . Using induction, we get that  $t^i(t-1)^j$  can be written as a sum of terms satisfying this condition.

The only question which remains is can we get  $j \leq i \leq 2j$ ? The only problem occurs if we get  $i < j$ .

Suppose that  $i > 2j$ . Then  $i \geq 2j + 1$ . Let  $i'$  and  $j'$  be the " $i$  and  $j$  of the next step". First suppose that  $j' = j$ ,  $i' = i - 1$ . Then, since  $i > 2j$ , so  $i \geq 2j + 1$ , we clearly have  $i' \geq 2j \geq j = j'$ . Thus  $i' \geq j'$ .

Now suppose that  $i' = i - 1$ ,  $j' = j + 1$ . Since  $i \geq 2j + 1 \geq j + 2$ , we get

$$i \geq j + 2 \implies i' \geq j + 1 = j'$$

Thus, once again  $i' \geq j'$ .

This means that if we stop on time after getting  $i \leq 2j$ ,  $i < j$  will not occur. Thus, for each  $k \in \mathbb{N}$ , we can write  $t^k(t-1)$  as a sum of terms  $t^i(t-1)^j$  satisfying  $j \leq i \leq 2j$ . Since each  $t^i(t-1)^j$ ,  $j \leq i \leq 2j$  is in the image of  $\varphi$ ,  $t^k(t-1)$  is in the image of  $\varphi$ . Finally, this means that the entire  $\tilde{p}(t)$  is in the image of  $\varphi$ . Since  $p(t) = \tilde{p}(t) + p(0)$ , it is also in the image of  $\varphi$ , as required.

## Result

7 of 7

(a) Hint:  $x^l y \rightsquigarrow t^{2l+3}$ .

$$(b) \ker \varphi = (x^3 - y^2 + xy)$$

If  $p(1) = p(0)$ , define  $\tilde{p}(t) = p(t) - p(0)$ . Then (why?)

$$\tilde{p}(t) = t(t-1)q(t) = \sum_{k=0}^n a_k t^{k+1}(t-1)$$

Now prove that each  $t^{k+1}(t-1)$  is in the image of  $\varphi$ .

6. a

Denote by  $\varphi$  the specified map.

Define

$$S = \{f(x) \in \mathbb{Z}[x] \mid f(x) \neq 0, f(\alpha) = 0\}$$

If  $S = \emptyset$ , then  $\ker \varphi = (0)$ , so it is a principal ideal.

If  $S \neq \emptyset$ , define

$$T = \{\deg f(x) \mid f(x) \in S\}$$

(with  $\deg f(x)$  we denote the degree of a polynomial  $f(x)$ ). Notice that  $T \subseteq \mathbb{N}$  ( $0 \notin T$  since nonconstant nonzero polynomials do not have  $\alpha$  as a root), so it has a minimal element!

Let  $g(x) \in S$  be one of the polynomials with the minimal degree. Suppose that it is not primitive. Let  $k$  be the greatest common divisor of its coefficients. Then we can write

$$g(x) = \pm kh(x),$$

where  $h(x)$  is primitive. Also,

$$0 = g(\alpha) = \pm kh(\alpha) \implies h(\alpha) = 0$$

Thus,  $h(x) \in S$ . Moreover,  $\deg h(x) = \deg g(x)$ , so  $h(x)$  also has the minimal degree in  $S$ !

Now we fix some primitive polynomial  $h(x) \in S$  with the minimal degree of all polynomials in  $S$ . Let  $f(x) \in \ker \varphi$ . We observe  $h(x)$  and  $f(x)$  as polynomials in  $\mathbb{Q}[x]$ , so we can divide  $f(x)$  by  $h(x)$ :

$$f(x) = h(x)q(x) + r(x),$$

where  $r(x) \in \mathbb{Q}[x]$  and  $\deg r(x) < \deg h(x)$ . Also,

$$0 = f(\alpha) = h(\alpha)q(\alpha) + r(\alpha) \implies r(\alpha) = 0$$

So, if  $r(x)$  is not a zero polynomial, we would have  $nr(x) \in S$ , where  $n$  is smallest common multiple of denominators of coefficients of  $r(x)$ . However,  $\deg nr(x) < \deg h(x)$ , which is impossible, since  $h(x)$  has the minimal degree. Thus, we must have that  $r(x) = 0$ , so

$$f(x) = h(x)q(x)$$

This means that  $h(x)$  is a primitive polynomial which divides  $f(x)$  in  $\mathbb{Q}[x]$ , so by **Theorem 12.3.6 (b)** we conclude that  $h(x)$  divides  $f(x)$  in  $\mathbb{Z}[x]$ . Thus,

$$f(x) = h(x)\tilde{q}(x)$$

for some  $\tilde{q}(x) \in \mathbb{Z}[x]$ , which means that

$$f(x) \in (h(x))$$

Thus,

$$\ker\varphi \subseteq (h(x))$$

For the other inclusion, let  $f(x) \in (h(x))$ . Then

$$f(x) = h(x)g(x)$$

for some  $g(x) \in \mathbb{Z}[x]$ . Moreover,

$$\varphi(f(x)) = f(\alpha) = h(\alpha)g(\alpha)$$

Thus,  $f(x) \in \ker\varphi$ , and

$$(h(x)) \subseteq \ker\varphi$$

Finally,

$$(h(x)) = \ker\varphi$$

## Result

3 of 3

Hint:  $\ker\varphi = (h(x))$ , where  $h(x)$  is some primitive polynomial such that  $h(\alpha) = 0$ , and  $h(x)$  has the minimal degree among all polynomials  $f(x)$  such that  $f(\alpha) = 0$ .

## Section 4

### 1. a

Consider the provided statement to factor the polynomial function in given field.

(a)

The factor  $x^9 - x$  in  $\mathbb{F}_3[x]$  is provided as below:

The monic irreducible polynomials of degree at most 3 over the field 3 are as,

$$\begin{aligned} &x, x+1, x-1, x^2+1, x^2+x-1, x^2-x-1 \\ &x^3-x+1, x^3-x-1, x^3+x^2-1, x^3-x^2+1, \\ &x^3+x^2+x+1, x^3+x^2+x-1, x^3+x^2-x+1, \\ &x^3-x^2+x+1, x^3-x^2+x-1, x^3-x^2-x-1 \end{aligned}$$

As the irreducible factors of a polynomial  $x^r - x$  over  $\mathbb{F}_3$  are the irreducible polynomials over  $\mathbb{F}_3$  whose degree divides  $r$ , therefore the factor of  $x^9 - x$  is provided as,

$$x^9 - x = x(x+1)(x-1)(x^2+1)(x^2+x-1)(x^2-x-1)$$



The factor  $x^9 - 1$  in  $\mathbb{F}_3[x]$  is provided as below:

The monic irreducible polynomials of degree at most 3 over the field 3 are as,

$$\begin{aligned} & x, x+1, x-1, x^2+1, x^2+x-1, x^2-x-1 \\ & x^3-x+1, x^3-x-1, x^3+x^2-1, x^3-x^2+1, \\ & x^3+x^2+x+1, x^3+x^2+x-1, x^3+x^2-x+1, \\ & x^3-x^2+x+1, x^3-x^2+x-1, x^3-x^2-x-1 \end{aligned}$$

As  $x^9 - 1$  can be also written in the form of  $(x^3)^3 - (1)^3$  therefore,

$$\begin{aligned} a^3 - b^3 &= (a-b)(a^2 + ab + b^2) \\ (x^3)^3 - (1)^3 &= (x^3 - 1)((x^3)^2 + x^3 + 1) \\ &= (x^3 - 1)(x^6 + x^3 + 1) \\ &= (x-1)(x^2 + x + 1)(x^6 + x^3 + 1) \end{aligned}$$

Therefore the factor of  $x^9 - x$  is,

$$x^9 - x = (x-1)(x^2 + x + 1)(x^6 + x^3 + 1)$$

(b)

The factor of the polynomial  $x^{16} - x$  in field  $\mathbb{F}_2[x]$  is provided by using the Maple software. As it is given that the field is 2  $\mathbb{F}_2$  then the value of modulus is 2.

$$\text{factor}(x^{16} - x) \text{ mod } 2 \quad x(x-1)(x^4+x^3+x^2+x+1)(x^2+x+1)(1-x+x^3-x^4+x^5-x^7+x^8)$$

Hence,

$$x^{16} - x = x(x-1)(x^2+x+1)(x^4+x^3+x^2+x+1)(x^8-x^7+x^5-x^4+x^3-x+1).$$

2. a

(a)

Suppose that it is not irreducible. Then there exists  $f(x) \in \mathbb{F}_7[x]$  which is not a unit, is not associated with  $x^2 + 1$ , and divides  $x^2 + 1$ . That is,

$$x^2 + 1 = f(x)g(x)$$

where  $g(x) \in \mathbb{F}_7[x]$ .

If  $f(x)$  is constant, then  $f(x) \neq 0$ , so it is a unit. If  $g(x)$  is constant, then  $g(x)$  is a unit, meaning that  $x^2 + 1$  and  $f(x)$  are associated. In both cases we arrived at a contradiction, thus both  $f(x)$  and  $g(x)$  are not constant. Since the product  $f(x)g(x)$  is a quadratic polynomial,  $f(x)$  and  $g(x)$  must be linear. Thus,

$$f(x) = ax + b, \quad g(x) = cx + d,$$

where  $a, b, c, d \in \mathbb{F}_7$ . Moreover,

$$f(a^{-1}b) = 0,$$

so  $x^2 + 1$  has a root  $a^{-1}b$ . However,

$$x = 0 \implies x^2 + 1 = 1 \neq 0$$

$$x = 1 \implies x^2 + 1 = 2 \neq 0$$

$$x = 2 \implies x^2 + 1 = 5 \neq 0$$

$$x = 3 \implies x^2 + 1 = 10 = 3 \neq 0$$

$$x = 4 \implies x^2 + 1 = 17 = 3 \neq 0$$

$$x = 5 \implies x^2 + 1 = 26 = 5 \neq 0$$

$$x = 6 \implies x^2 + 1 = 37 = 2 \neq 0$$

Thus,  $x^2 + 1$  has no roots in  $\mathbb{F}_7$ . Contradiction! Thus,  $x^2 + 1$  must be irreducible.

**(b)**

Similarly to **(a)**, we assume that it is not irreducible, so

$$x^3 - 9 = f(x)g(x)$$

where  $f(x)$  and  $g(x)$  are nonconstant. Since their product is of degree 3, one of them must be of degree 1, while the other must be of degree 2. Without loss of generality assume that

$$f(x) = ax + b, \quad g(x) = cx^2 + dx + e,$$

with  $a, b, c, d, e \in \mathbb{F}_{31}$ . Again,  $f(a^{-1}b) = 0$ , so  $x^3 - 9$  must have a root in  $\mathbb{F}_{31}$ . As in **(a)**, we check manually for roots and check that  $x^3 - 9$  has no roots in  $\mathbb{F}_{31}$ . Hence,  $x^3 - 9$  is irreducible in  $\mathbb{F}_{31}$ .

## Result

3 of 3

Hint: these polynomials have no roots in respective fields. If they were not irreducible, they would have to have roots (why?).

3. a

$$\text{Let } f(x) = x^4 + 6x^3 + 9x + 3.$$

Notice that 3 is a prime number such that

- 3 does not divide 1 (the leading coefficient of  $f(x)$ ).
- 3 divides all other coefficients.
- $9 = 3^2$  does not divide 3 (the constant coefficient of  $f(x)$ ).

By the **Eisenstein Criterion (Proposition 12.4.6)**, we conclude that  $f(x)$  is irreducible in  $\mathbb{Q}[x]$ .

Since  $\mathbb{Q}$  is a field, by **Proposition 12.2.5 (b)** we conclude that  $\mathbb{Q}[x]$  is a Euclidean domain. By **Proposition 12.2.7**, we now conclude that  $\mathbb{Q}[x]$  is a principal ideal domain. By **Corollary 12.2.9 (c)** this means that  $(f(x))$  is a maximal ideal in  $\mathbb{Q}[x]$ .

## Result

2 of 2

Hint:  $f(x)$  is irreducible in  $\mathbb{Q}[x]$ .

4. a

### Modulo 2

This polynomial modulo 2 is

$$f(x) = x^5 + x^3 + x + 1$$

Furthermore,  $f(1) = 4 = 0$ , so  $x - 1 = x + 1$  divides it. So,

$$x^5 + x^3 + x + 1 = (x + 1)(ax^4 + bx^3 + cx^2 + dx + e)$$

where  $a, b, c, d, e \in \mathbb{F}_2$ . Furthermore,

$$x^5 + x^3 + x + 1 = ax^5 + (a + b)x^4 + (b + c)x^3 + (c + d)x^2 + (d + e)x + e$$

Thus,

$$\begin{aligned}
a &= 1 \\
a + b &= 0 \\
b + c &= 1 \\
c + d &= 0 \\
d + e &= 1 \\
e &= 1
\end{aligned}$$

From this we get

$$a = 1, \quad b = 1, \quad c = 0, \quad d = 0, \quad e = 1$$

Therefore,

$$x^5 + x^3 + x + 1 = (x + 1)(x^4 + x^3 + 1)$$

Looking at (12.4.4) in the book (page 373), we see that  $x^4 + x^3 + 1$  is irreducible. Furthermore,  $x + 1$  is also irreducible. Thus, we cannot factor this further, so we conclude that

$$x^5 + x^3 + x + 1 = (x + 1)(x^4 + x^3 + 1)$$

### Modulo 3

This polynomial modulo 3 is

$$f(x) = x^5 + 2x^4 + 2$$

Notice that

$$f(2) = 66 = 0,$$

so  $x - 2 = x + 1$  divides  $f(x)$ . Again, we try to find  $a, b, c, d, e \in \mathbb{F}_3$  such that

$$x^5 + 2x^4 + 2 = (x + 1)(ax^4 + bx^3 + cx^2 + dx + e)$$

Since

$$(x + 1)(ax^4 + bx^3 + cx^2 + dx + e) = ax^5 + (a + b)x^4 + (b + c)x^3 + (c + d)x^2 + (d + e)x + e,$$

we get the system of equations

$$\begin{aligned}
a &= 1 \\
a + b &= 2 \\
b + c &= 0 \\
c + d &= 0 \\
d + e &= 0 \\
e &= 2
\end{aligned}$$

From this we get

$$a = 1, \quad b = 1, \quad c = 2, \quad d = 1, \quad e = 2$$

Therefore,

$$f(x) = x^5 + 2x^4 + 2 = (x+1)(x^4 + x^3 + 2x^2 + x + 2)$$

Let  $g(x) = x^4 + x^3 + 2x^2 + x + 2$ . Notice that

$$g(2) = 36 = 0$$

Thus,  $x - 2 = x + 1$  divides  $g(x)$ , so we need to find  $a, b, c, d \in \mathbb{F}_3$  such that

$$x^4 + x^3 + 2x^2 + x + 2 = (x+1)(ax^3 + bx^2 + cx + d)$$

Since

$$(x+1)(ax^3 + bx^2 + cx + d) = ax^4 + (a+b)x^3 + (b+c)x^2 + (c+d)x + d,$$

we get a system of equations

$$\begin{aligned} a &= 1 \\ a+b &= 1 \\ b+c &= 2 \\ c+d &= 1 \\ d &= 2 \end{aligned}$$

From this we get

$$a = 1, \quad b = 0, \quad c = 2, \quad d = 2$$

Therefore,

$$g(x) = x^4 + x^3 + 2x^2 + x + 2 = (x+1)(x^3 + 2x + 2)$$

Furthermore,

$$f(x) = (x+1)g(x) = (x+1)^2(x^3 + 2x + 2)$$

Now suppose that we can factor  $x^3 + 2x + 2$  using nonconstant polynomials; that is, we can write

$$x^3 + 2x + 2 = h_1(x)h_2(x)$$

where  $h_1(x)$  and  $h_2(x)$  are nonconstant. Then one of them is of degree 1, while the other is of degree 2. Without loss of generality assume that

$$h_1(x) = ax + b$$

Then  $h_1(a^{-1}b) = 0$ , so  $h_1$  has a root. But this means that  $x^3 + 2x + 2$  must also have a root in  $\mathbb{F}_3$ . However,

$$x = 0 \implies x^3 + 2x + 2 = 2 \neq 0$$

$$x = 1 \implies x^3 + 2x + 2 = 5 = 2 \neq 0$$

$$x = 2 \implies x^3 + 2x + 2 = 16 = 1 \neq 0$$

Thus,  $x^3 + 2x + 2$  has no roots in  $\mathbb{F}_3$ . Contradiction. Thus, we cannot factor  $x^3 + 2x + 2$  further (this also means that  $x^3 + 2x + 2$  is irreducible in  $\mathbb{F}_3[x]$ ).

Finally, we now have that

$$\boxed{x^5 + 2x^4 + 2 = (x+1)^2(x^3 + 2x + 2)}$$

Let  $f(x) = x^5 + 2x^4 + 3x^3 + 3x + 5$ . Notice that  $f(-1) = 0$ , so  $x - (-1) = x + 1$  divides  $f(x)$ . So, we want to find  $a, b, c, d, e \in \mathbb{Q}$  such that

$$x^5 + 2x^4 + 3x^3 + 3x + 5 = (x + 1)(ax^4 + bx^3 + cx^2 + dx + e)$$

Since

$$(x + 1)(ax^4 + bx^3 + cx^2 + dx + e) = ax^5 + (a + b)x^4 + (b + c)x^3 + (c + d)x^2 + (d + e)x + e$$

This yields the system of equations

$$\begin{aligned} a &= 1 \\ a + b &= 2 \\ b + c &= 3 \\ c + d &= 0 \\ d + e &= 3 \\ e &= 5 \end{aligned}$$

From this we get

$$a = 1, \quad b = 1, \quad c = 2, \quad d = -2, \quad e = 5$$

Thus,

$$f(x) = (x + 1)(x^4 + x^3 + 2x^2 - 2x + 5)$$

Now we will prove that  $x^4 + x^3 + 2x^2 - 2x + 5$  is irreducible in  $\mathbb{Q}[x]$ . Its residue in  $\mathbb{F}_2[x]$  is

$$x^4 + x^3 + 1,$$

which is an irreducible polynomial in  $\mathbb{F}_2[x]$  (see (12.4.4) in the book). Thus, by **Proposition 12.4.3** we conclude that  $x^4 + x^3 + 2x^2 - 2x + 5$  is irreducible in  $\mathbb{Q}[x]$ . Therefore, we cannot factor  $f(x)$  further, meaning that

$$f(x) = (x + 1)(x^4 + x^3 + 2x^2 - 2x + 5)$$

## Result

5 of 5

$$\text{(a)} \quad (x + 1)(x^4 + x^3 + 1)$$

$$\text{(b)} \quad (x + 1)^2(x^3 + 2x + 2)$$

$$\text{(c)} \quad (x + 1)(x^4 + x^3 + 2x^2 - 2x + 5)$$

5. a



Recall **the Eisenstein's Criterion**: A polynomial  $f(x) \in \mathbb{Q}[x]$  is irreducible in  $\mathbb{Q}[x]$  if there exists a prime integer  $p$  such that the following holds:

1.  $p$  does not divide  $a_n$  (the leading coefficient),
2.  $p$  divides all other coefficients,
3.  $p^2$  does not divide  $a_0$  (the constant coefficient).

## Step 2

2 of 6

### (a)

Notice that 3 is a prime integer such that

1. 3 does not divide 1 (the leading coefficient),
2. 3 divides all other coefficients,
3.  $9 = 3^2$  does not divide 213 (the constant coefficient).

Thus, this polynomial is irreducible.

### (b)

Here the Eisenstein's Criterion will not work, but we can still prove that this polynomial is irreducible.

Suppose that it is not irreducible, so there exists  $f(x) \in \mathbb{Q}[x]$  which is not a unit and is not associated with  $8x^3 - 6x + 1$ . Thus,

$$8x^3 - 6x + 1 = f(x)g(x),$$

for some  $g(x) \in \mathbb{Q}[x]$ .

If  $f(x)$  is constant, then it must be  $f(x) \neq 0$ , so it is a unit. Thus,  $f(x)$  is not constant.

Similarly, if  $g(x)$  is constant, then it is a unit, so  $f(x)$  and  $8x^3 - 6x + 1$  are associated.

Thus,  $f(x)$  and  $g(x)$  are not constant. Since their product is of degree 3, one of them must be of degree 1, while the other one must be of degree 2. Suppose that

$$f(x) = ax + b,$$

for some  $a, b \in \mathbb{Q}$ ,  $a \neq 0$ . Then  $f(-b/a) = 0$ , so  $f(x)$  has a root in  $\mathbb{Q}$ . This means that  $8x^3 - 6x + 1$  must also have a rational root!

We get the same conclusion if we assume that  $g(x)$  is linear.

Now we want to prove that  $8x^3 - 6x + 1$  does not have a rational root. By **the Rational root theorem**, if  $\frac{a}{b}$  is a rational root, we must have that  $a$  divides 1 and  $b$  divides 8. Thus, all possibilities are

$$\pm 1, \quad \pm \frac{1}{2}, \quad \pm \frac{1}{4}, \quad \pm \frac{1}{8}$$

Manually checking all of them, we see that none are roots of  $8x^3 - 6x + 1$ , so this polynomial has no rational roots. Moreover, we now conclude that  $f(x)$  and  $g(x)$ , both nonconstant, such that  $8x^3 - 6x + 1 = f(x)g(x)$  do not exist. Thus,  $8x^3 - 6x + 1$  is irreducible in  $\mathbb{Q}[x]$ .



(c)

This is solved the same way as (b). Here the possibilities for rational roots are

$$\pm 1$$

Manually checking them, we see that none are roots of  $x^3 + 6x^2 + 1$ , as in (b), we conclude that this polynomial is irreducible in  $\mathbb{Q}[x]$ .

### Step 5

5 of 6

(d)

Notice that 3 is a prime integer such that

1. 3 does not divide 1 (the leading coefficient),
2. 3 divides all other coefficients,
3.  $9 = 3^2$  does not divide 3 (the constant coefficient).

Thus, this polynomial is irreducible.

### Result

6 of 6

All are irreducible. For (a) and (d) you can use the Eisenstein's Criterion, while (b) and (c) are proven a bit more directly.

6. a

$\mathbb{Q}[x]$

Recall **the Eisenstein's Criterion**: A polynomial  $f(x) \in \mathbb{Q}[x]$  is irreducible in  $\mathbb{Q}[x]$  if there exists a prime integer  $p$  such that the following holds:

1.  $p$  does not divide  $a_n$  (the leading coefficient),
2.  $p$  divides all other coefficients,
3.  $p^2$  does not divide  $a_0$  (the constant coefficient).

Notice that 5 is a prime integer such that

1. 5 does not divide 1 (the leading coefficient),
2. 5 divides all other coefficients,
3.  $25 = 5^2$  does not divide 5 (the constant coefficient).

Thus, this polynomial is irreducible.

### $\mathbb{F}_2[x]$

Here we can first conclude that

$$x^5 + 5x + 5 = x^5 + x + 1$$

Now suppose that

$$x^5 + x + 1 = f(x)g(x),$$

where  $f(x)$  and  $g(x)$  are nonconstant, and one of them is irreducible (we now that if  $x^5 + x + 1$  is not irreducible, then it has at least one irreducible factor). Moreover, the degree of  $f(x)$  and  $g(x)$  are 4 or less, so we can look at the list of irreducible polynomials in  $\mathbb{F}_2[x]$  provided in the book (page 373). From this, we can see that

$$x^5 + x + 1 = (x^3 + x^2 + 1)(x^2 + x + 1)$$

Furthermore,  $x^3 + x^2 + 1$  and  $x^2 + x + 1$  are both irreducible (look at the list!), so we cannot factor this further.

### Result

In  $\mathbb{Q}[x]$ ,  $x^5 + 5x + 5$  is already irreducible.

In  $\mathbb{F}_2[x]$ ,

$$x^5 + 5x + 5 = x^5 + x + 1 = (x^3 + x^2 + 1)(x^2 + x + 1)$$

7. a

### $\mathbb{F}_2[x]$

Here  $x^3 + x + 1$  is already irreducible (see page 373 in the book), so it is already factored.

### $\mathbb{F}_3[x]$

Notice that 1 is a root of  $x^3 + x + 1$  in  $\mathbb{F}_3[x]$ . Thus,  $x - 1 = x + 2$  divides  $x^3 + x + 1$ . Now we must find a polynomial  $g(x) \in \mathbb{F}_3[x]$  such that

$$x^3 + x + 1 = (x + 2)g(x)$$

(such exists because  $x + 2$  divides  $x^3 + x + 1$ ). Moreover,  $g(x)$  is of degree 2 (because  $\deg((x + 2)g(x)) = 1 + \deg g(x)$  and  $\deg(x^3 + x + 1) = 3$ , so  $\deg((x + 2)g(x)) = 3$ , which implies  $\deg g(x) = 2$ ). Thus,

$$g(x) = ax^2 + bx + c,$$

for some  $a, b, c \in \mathbb{F}_3$ ,  $a \neq 0$ . So,

$$x^3 + x + 1 = (x + 2)(ax^2 + bx + c) = ax^3 + (2a + b)x^2 + (2b + c)x + 2c$$

This yields the system of equations

$$\begin{aligned} a &= 1 \\ 2a + b &= 0 \\ 2b + c &= 1 \\ 2c &= 1 \end{aligned}$$

Thus,  $a = 1$ ,  $b = -2 = 1$ ,  $c = 1 - 2 = -1 = 2$ , and

$$g(x) = x^2 + x + 2$$

Finally,

$$x^3 + x + 1 = (x + 2)(x^2 + x + 2)$$

Furthermore,  $x + 2$  is clearly irreducible, while  $x^2 + x + 1$  is irreducible by (12.4.5) in the book. Thus, this is the final factorization:

$$x^3 + x + 1 = (x + 2)(x^2 + x + 2)$$

### $\mathbb{F}_5[x]$

We will prove that this polynomial is irreducible in  $\mathbb{F}_5[x]$ . Suppose that  $f(x) \in \mathbb{F}_5[x]$  is not a unit and is not associated with  $x^3 + x + 1$ , and it divides  $x^3 + x + 1$ . Thus, there exists  $g(x) \in \mathbb{F}_5[x]$  such that

$$x^3 + x + 1 = f(x)g(x) \quad (1)$$

If  $f(x)$  is constant, then  $f(x) \neq 0$ , so it is a unit.

If  $g(x)$  is constant, then  $g(x)$  is a unit, and  $f(x)$  and  $x^3 + x + 1$  are associated.

Thus, neither  $f(x)$  nor  $g(x)$  are constant. Moreover, the sum of their degrees must be 3 (because  $x^3 + x + 1$  is of degree 3). Therefore, one of them must be of degree 1. Suppose that  $f(x) = ax + b$ ,  $a \neq 0$ . Then  $f(-b/a) = 0$ , so it has a root, and  $x^3 + x + 1$  must also have a root at  $x = -b/a$ . We conclude the same when we assume that  $g(x)$  is of degree 1.

Now we prove that  $x^3 + x + 1$  does not have a root:

$$\begin{aligned} x = 0 &\implies x^3 + x + 1 = 1 \neq 0 \\ x = 1 &\implies x^3 + x + 1 = 3 \neq 0 \\ x = 2 &\implies x^3 + x + 1 = 11 = 1 \neq 0 \\ x = 3 &\implies x^3 + x + 1 = 31 = 1 \neq 0 \\ x = 4 &\implies x^3 + x + 1 = 69 = 4 \neq 0 \end{aligned}$$

So,  $x^3 + x + 1$  does not have a root. This means that the factorization (1) cannot hold.

Thus,  $x^3 + x + 1$  is irreducible, and we cannot factor it further (in  $\mathbb{F}_5[x]$ ).

### Result

3 of 3

$\mathbb{F}_2[x]$ : This polynomial is irreducible, so we cannot factor it further.

$$\mathbb{F}_3[x]: x^3 + x + 1 = (x + 2)(x^2 + x + 2)$$

$\mathbb{F}_5[x]$ : This polynomial is irreducible, so we cannot factor it further.

8. a

A good idea is to put  $t = x^2$ , so we get a polynomial in  $F[t]$ :

$$t^2 + bt + c$$

We try to factor it:

$$t^2 + bt + c = f(t)g(t),$$

for  $f(t), g(t) \in F[t]$  nonconstant. Thus, both  $f(t)$  and  $g(t)$  must be of degree 2, leading to

$$t^2 + bt + c = (m_1t + n_1)(m_2t + n_2),$$

where  $m_1, n_1, m_2, n_2 \in F$ . After returning  $t = x^2$ , we get

$$x^4 + bx^2 + c = (m_1x^2 + n_1)(m_2x^2 + n_2)$$

So, for  $x^4 + 4x^2 + 4$ , setting  $t = x^2$  we get

$$t^2 + 4t + 4$$

Also,

$$t^2 + 4t + 4 = (t + 2)^2$$

Thus,

$$x^4 + 4x^2 + 4 = (x^2 + 2)^2$$

For  $x^4 + 3x^2 + 4$ , setting  $t = x^2$  yields

$$t^2 + 3t + 4$$

Now we do not know if we can factor it further; it depends which field  $F$  we pick.

## Result

Hint: set  $t = x^2$ .

9. a

If  $n = 0$ , then the polynomial is  $-p$ , which is clearly not irreducible (since it is a unit). For  $n < 0$ , this expression is not even a polynomial. So, let  $n > 0$ , and let  $p$  be any prime integer. Then

- $p$  does not divide 1 (the leading coefficient).
- $p$  divides all other coefficients.
- $p^2$  does not divide  $-p$  (the constant coefficient).

Thus, by **the Eisenstein Criterion (Proposition 12.4.6)**, we conclude that  $x^n - p$  is irreducible.

## Result

2 of 2

Any prime integer  $p$  and any positive integer  $n$ .

10. a

(a)

To factor the polynomial  $(x^2 + 2351x + 125)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^2 + 2351x + 125$$

Since, the polynomial  $(x^2 + 2351x + 125)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^2 + 2351x + 125) \bmod(5) \\ (x + 1) x$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_5[x]$  is,

$$x^2 + 2351x + 125 \equiv x(x+1) \pmod{5}$$

Since, the factors of the polynomial in  $\mathbb{F}_5[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^2 + 2351x + 125 \equiv x(x+1) \pmod{5}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is  $\boxed{x^2 + 2351x + 125 = x(x+1)}$ .

---

(b)

To factor the polynomial  $(x^3 + 2x^2 + 3x + 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^3 + 2x^2 + 3x + 1$$

Since, the polynomial  $(x^3 + 2x^2 + 3x + 1)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^3 + 2x^2 + 3x + 1) \bmod(5) \\ (x + 2) (x^2 + 3)$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_5[x]$  is,

$$x^3 + 2x^2 + 3x + 1 \equiv (x+2)(x^2+3) \pmod{5}$$

Since, the factors of the polynomial in  $\mathbb{F}_5[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^3 + 2x^2 + 3x + 1 \equiv (x+2)(x^2+3) \pmod{5}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$\boxed{x^3 + 2x^2 + 3x + 1 = (x+2)(x^2+3)}.$$



(c)

To factor the polynomial  $(x^4 + 2x^3 + 2x^2 + 2x + 2)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^4 + 2x^3 + 2x^2 + 2x + 2$$

Since, the polynomial  $(x^4 + 2x^3 + 2x^2 + 2x + 2)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^4 + 2x^3 + 2x^2 + 2x + 2) \bmod(3) \\ (x + 2)(x^3 + 2x + 1)$$

31

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_3[x]$  is,

$$x^4 + 2x^3 + 2x^2 + 2x + 2 \equiv (x + 2)(x^3 + 2x + 1) \pmod{3}$$

Since, the factors of the polynomial in  $\mathbb{F}_3[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^4 + 2x^3 + 2x^2 + 2x + 2 \equiv (x + 2)(x^3 + 2x + 1) \pmod{3}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$x^4 + 2x^3 + 2x^2 + 2x + 2 = (x + 2)(x^3 + 2x + 1).$$

(d)

To factor the polynomial  $(x^4 + 2x^3 + 3x^2 + 2x + 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^4 + 2x^3 + 3x^2 + 2x + 1$$

Since, the polynomial  $(x^4 + 2x^3 + 3x^2 + 2x + 1)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^4 + 2x^3 + 3x^2 + 2x + 1) \bmod(2) \\ (x^2 + x + 1)^2$$

31

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^4 + 2x^3 + 3x^2 + 2x + 1 \equiv (x^2 + x + 1)^2 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^4 + 2x^3 + 3x^2 + 2x + 1 \equiv (x^2 + x + 1)^2 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$x^4 + 2x^3 + 3x^2 + 2x + 1 = (x^2 + x + 1)^2.$$



(e)

To factor the polynomial  $(x^4 + 2x^3 + x^2 + 2x + 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^4 + 2x^3 + x^2 + 2x + 1$$

Since, the polynomial  $(x^4 + 2x^3 + x^2 + 2x + 1)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^4 + 2x^3 + x^2 + 2x + 1) \bmod(2) \\ (x^2 + x + 1)^2$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^4 + 2x^3 + x^2 + 2x + 1 \equiv (x^2 + x + 1)^2 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^4 + 2x^3 + x^2 + 2x + 1 \equiv (x^2 + x + 1)^2 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$\boxed{x^4 + 2x^3 + x^2 + 2x + 1 = (x^2 + x + 1)^2}.$$

(f)

To factor the polynomial  $(x^4 + 2x^2 + x + 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^4 + 2x^2 + x + 1$$

Since, the polynomial  $(x^4 + 2x^2 + x + 1)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^4 + 2x^2 + 2x + 1) \bmod(2) \\ (x + 1)^4$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^4 + 2x^2 + x + 1 \equiv (x + 1)^4 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^4 + 2x^2 + x + 1 \equiv (x + 1)^4 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is  $\boxed{x^4 + 2x^2 + x + 1 = (x + 1)^4}$ .

(g)

To factor the polynomial  $(x^8 + x^6 + x^4 + x^2 + 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^8 + x^6 + x^4 + x^2 + 1$$

Since, the polynomial  $(x^8 + x^6 + x^4 + x^2 + 1)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^8 + x^6 + x^4 + x^2 + 1) \bmod(2) \\ (x^4 + x^3 + x^2 + x + 1)^2$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^8 + x^6 + x^4 + x^2 + 1 \equiv (x^4 + x^3 + x^2 + x + 1)^2 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^8 + x^6 + x^4 + x^2 + 1 \equiv (x^4 + x^3 + x^2 + x + 1)^2 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$\boxed{x^8 + x^6 + x^4 + x^2 + 1 = (x^4 + x^3 + x^2 + x + 1)^2}.$$

(h)

To factor the polynomial  $(x^6 - 2x^5 - 3x^2 + 9x - 3)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^6 - 2x^5 - 3x^2 + 9x - 3$$

Since, the polynomial  $(x^6 - 2x^5 - 3x^2 + 9x - 3)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^6 - 2x^5 - 3x^2 + 9x - 3) \bmod(2) \\ (x^5 + x^4 + x^3 + x^2 + 1)(x + 1)$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^6 - 2x^5 - 3x^2 + 9x - 3 \equiv (x^5 + x^4 + x^3 + x^2 + 1)(x + 1) \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^6 - 2x^5 - 3x^2 + 9x - 3 \equiv (x^5 + x^4 + x^3 + x^2 + 1)(x + 1) \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$\boxed{x^6 - 2x^5 - 3x^2 + 9x - 3 = (x^5 + x^4 + x^3 + x^2 + 1)(x + 1)}.$$

(i)

To factor the polynomial  $(x^4 + x^2 + 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^4 + x^2 + 1$$

Since, the polynomial  $(x^4 + x^2 + 1)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^4 + x^2 + 1) \bmod(2) \\ (x^2 + x + 1)^2$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^4 + x^2 + 1 \equiv (x^2 + x + 1)^2 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^4 + x^2 + 1 \equiv (x^2 + x + 1)^2 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is  $x^4 + x^2 + 1 = (x^2 + x + 1)^2$ .

(j)

To factor the polynomial  $(3x^5 + 6x^4 + 9x^3 + 3x^2 - 1)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$3x^5 + 6x^4 + 9x^3 + 3x^2 - 1$$

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(3x^5 + 6x^4 + 9x^3 + 3x^2 - 1) \bmod(2) \\ (x^2 + x + 1)(x + 1)^3$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$3x^5 + 6x^4 + 9x^3 + 3x^2 - 1 \equiv (x^2 + x + 1)(x + 1)^3 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$3x^5 + 6x^4 + 9x^3 + 3x^2 - 1 \equiv (x^2 + x + 1)(x + 1)^3 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$3x^5 + 6x^4 + 9x^3 + 3x^2 - 1 = (x^2 + x + 1)(x + 1)^3.$$

(k)

To factor the polynomial  $(x^5 + x^4 + x^2 + x + 2)$  in  $\mathbb{Q}[x]$ ,

Consider the following polynomial

$$x^5 + x^4 + x^2 + x + 2$$

Since, the polynomial  $(x^5 + x^4 + x^2 + x + 2)$  is the monic integer polynomial.

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

Then, the factors of the polynomial in  $\mathbb{F}_p[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Use MAPLE to factor the above polynomial in  $\mathbb{Q}[x]$ .

$$\text{Factor}(x^5 + x^4 + x^2 + x + 2) \bmod(2)$$

$$(x^2 + x + 1)(x + 1)^2 x$$

∴

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$x^5 + x^4 + x^2 + x + 2 \equiv x(x^2 + x + 1)(x + 1)^2 \pmod{2}$$

Since, the factors of the polynomial in  $\mathbb{F}_2[x]$  will be the factors in  $\mathbb{Q}[x]$  under the same modulo.

Then, the factorization of the polynomial in  $\mathbb{Q}[x]$  is,

$$x^5 + x^4 + x^2 + x + 2 \equiv x(x^2 + x + 1)(x + 1)^2 \pmod{2}$$

Hence, the required factorization of the polynomial in  $\mathbb{Q}[x]$  is

$$x^5 + x^4 + x^2 + x + 2 = x(x^2 + x + 1)(x + 1)^2.$$

11. a

We first put all integers  $2 \leq n \leq 99$  into a table:

2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	32	33	34	35	36
37	38	39	40	41	42	43
44	45	46	47	48	49	50
51	52	53	54	55	56	57
58	59	60	61	62	63	64
65	66	67	68	69	70	71
72	73	74	75	76	77	78
79	80	81	82	83	84	85
86	87	88	89	90	91	92
93	94	95	96	97	98	99

Now we mark 2 as prime, and remove from the list of candidates (mark with some other color) all multiples of 2:

2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	32	33	34	35	36
37	38	39	40	41	42	43
44	45	46	47	48	49	50
51	52	53	54	55	56	57
58	59	60	61	62	63	64
65	66	67	68	69	70	71
72	73	74	75	76	77	78
79	80	81	82	83	84	85
86	87	88	89	90	91	92
93	94	95	96	97	98	99

Now we look at the smallest integer which is not marked. We mark 3 as prime, and remove all multiples of 3 from the list of candidates:

2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	32	33	34	35	36
37	38	39	40	41	42	43
44	45	46	47	48	49	50
51	52	53	54	55	56	57
58	59	60	61	62	63	64
65	66	67	68	69	70	71
72	73	74	75	76	77	78
79	80	81	82	83	84	85
86	87	88	89	90	91	92
93	94	95	96	97	98	99

Now we look at the smallest integer which is not marked. We mark 5 as prime, and remove all multiples of 5 from the list of candidates:

2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	32	33	34	35	36
37	38	39	40	41	42	43
44	45	46	47	48	49	50
51	52	53	54	55	56	57
58	59	60	61	62	63	64
65	66	67	68	69	70	71
72	73	74	75	76	77	78
79	80	81	82	83	84	85
86	87	88	89	90	91	92
93	94	95	96	97	98	99



Now we look at the smallest integer which is not marked. We mark 7 as prime, and remove all multiples of 7 from the list of candidates:

2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	32	33	34	35	36
37	38	39	40	41	42	43
44	45	46	47	48	49	50
51	52	53	54	55	56	57
58	59	60	61	62	63	64
65	66	67	68	69	70	71
72	73	74	75	76	77	78
79	80	81	82	83	84	85
86	87	88	89	90	91	92
93	94	95	96	97	98	99

Similarly, we now get that all other unmarked numbers are prime:

2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	32	33	34	35	36
37	38	39	40	41	42	43
44	45	46	47	48	49	50
51	52	53	54	55	56	57
58	59	60	61	62	63	64
65	66	67	68	69	70	71
72	73	74	75	76	77	78
79	80	81	82	83	84	85
86	87	88	89	90	91	92
93	94	95	96	97	98	99

## Result

7 of 7

This algorithm removes nonprime integers from the list fast at the start, but it slows down afterwards. However, it is still much faster than manually checking for each integer whether it is prime or not.

12. a



(a)

All monic polynomials of degree 3 in  $\mathbb{F}_3$  are of the form

$$f(x) = x^3 + ax^2 + bx + c,$$

for some  $a, b, c \in \mathbb{F}_3$ .

Suppose that  $f(x)$  is not irreducible. Then there exists  $g(x) \in \mathbb{F}_3[x]$  which is not a unit and is not associated with  $f(x)$ , and it divides  $f(x)$ :

$$f(x) = g(x)h(x),$$

for some  $h(x) \in \mathbb{F}_3[x]$ .

If  $g(x)$  is constant, then  $g(x) \neq 0$ , and it is a unit.

If  $h(x)$  is constant, then it is a unit, so  $f(x)$  and  $g(x)$  are associated.

Thus, neither  $g(x)$  nor  $h(x)$  are constants. This means that one of them is of degree 1, while the other is of degree 2. If  $g(x) = dx + e$ , then  $g(-ed^{-1}) = 0$ , so  $g(x)$  has a root. But this means that  $f(x)$  also has a root!

Similarly, if  $h(x)$  is of degree 1,  $f(x)$  again has a root.

So, if  $f(x)$  is not irreducible, it has a root in  $\mathbb{F}_3[x]$ . The converse is also true; if it has a root  $x_0$ , then  $x - x_0$  divides it, which is not a unit nor is it associated with  $f(x)$ , so  $f(x)$  is not irreducible.

If  $c = 0$ , then  $f(x) = x^3 + ax^2 + bx$ , so  $f(0) = 0$ , which is not irreducible. Now we list all candidates for  $f(x)$  and manually check if they have roots in  $\mathbb{F}_3$  (note that 0 is clearly not a root):

$$\begin{aligned} f(x) = x^3 + 1 &\implies f(2) = 8 + 1 = 9 = 0 \implies 2 \text{ is a root!} \\ f(x) = x^3 + 2 &\implies f(1) = 1 + 2 = 3 = 0 \implies 1 \text{ is a root!} \\ f(x) = x^3 + x + 1 &\implies f(1) = 1 + 1 + 1 = 3 = 0 \implies 1 \text{ is a root} \\ f(x) = x^3 + x + 2 &\implies f(2) = 8 + 2 + 2 = 12 = 0 \implies 2 \text{ is a root!} \\ f(x) = x^3 + 2x + 1 &\implies f(1) = 4 \neq 0, f(2) = 13 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + 2x + 2 &\implies f(1) = 5 \neq 0, f(2) = 14 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + x^2 + 1 &\implies f(1) = 3 = 0 \implies 1 \text{ is a root!} \\ f(x) = x^3 + x^2 + 2 &\implies f(1) = 4 \neq 0, f(2) = 16 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + x^2 + x + 1 &\implies f(2) = 15 = 0 \implies 2 \text{ is a root!} \\ f(x) = x^3 + x^2 + x + 2 &\implies f(1) = 5 \neq 0, f(2) = 16 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + x^2 + 2x + 1 &\implies f(1) = 5 \neq 0, f(2) = 17 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + x^2 + 2x + 2 &\implies f(2) = 18 = 0 \implies 2 \text{ is a root!} \\ f(x) = x^3 + 2x^2 + 1 &\implies f(1) = 4 \neq 0, f(2) = 17 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + 2x^2 + 2 &\implies f(2) = 18 = 0 \implies 2 \text{ is a root!} \\ f(x) = x^3 + 2x^2 + x + 1 &\implies f(1) = 5 \neq 0, f(2) = 19 \neq 0 \implies \text{no roots!} \\ f(x) = x^3 + 2x^2 + x + 2 &\implies f(1) = 6 = 0 \implies 1 \text{ is a root!} \\ f(x) = x^3 + 2x^2 + 2x + 1 &\implies f(1) = 6 = 0 \implies 1 \text{ is a root!} \\ f(x) = x^3 + 2x^2 + 2x + 2 &\implies f(1) = 7 \neq 0, f(2) = 22 \neq 0 \implies \text{no roots!} \end{aligned}$$

Thus, all monic irreducible polynomials of degree 3 are

$$\begin{aligned} & x^3 + 2x + 1, \quad x^3 + 2x + 2, \quad x^3 + x^2 + 2, \quad x^3 + x^2 + x + 2, \\ & x^3 + x^2 + 2x + 1, \quad x^3 + 2x^2 + 1, \quad x^3 + 2x^2 + x + 1, \quad x^3 + 2x^2 + 2x + 2 \end{aligned}$$

### Step 3

3 of 6

#### (b)

This is similar to (a), except that we conclude that  $g(x)$  and  $h(x)$  such that  $f(x) = g(x)h(x)$  must be of degree 1. So,

$$f(x) = x^2 + ax + b,$$

where  $a, b \in \mathbb{F}_5$ . Again,  $b \neq 0$ . Manually checking, we see that the required polynomials are

$$\begin{aligned} & x^2 + x + 1, \quad x^2 + 4x + 1, \quad x^2 + 2, \quad x^2 + x + 2, \quad x^2 + 4x + 2, \\ & x^2 + 3, \quad x^2 + 2x + 3, \quad x^2 + 3x + 3, \quad x^2 + 2x + 4, \quad x^2 + 3x + 4 \end{aligned}$$

#### (c)

We will first find the number of monic irreducible polynomials of degree 3. Since  $\mathbb{F}_5[x]$  is a unique factorization domain, each monic  $f(x) \in \mathbb{F}_5[x]$  can be factored uniquely (up to the ordering of the terms) as

$$f(x) = p_1(x) \cdots p_n(x),$$

where  $p_1(x), \dots, p_n(x) \in \mathbb{F}_5[x]$  are monic and irreducible.

It is easier to find which polynomials are not irreducible. They factor as either

$$f(x) = p_1(x)p_2(x)p_3(x),$$

where  $p_1(x), p_2(x), p_3(x)$  are irreducible monic polynomials in  $\mathbb{F}_5[x]$  of degree 1, or

$$f(x) = p(x)q(x),$$

where  $p(x)$  is of degree 1, and  $q(x)$  is of degree 2, both irreducible and monic.

Since each  $x + a$ ,  $a \in \mathbb{F}_5$ , is clearly monic and irreducible, there are 5 such polynomials.

From (b), there are 10 monic irreducible polynomials of degree 2.

So, if  $f(x)$  is not irreducible and it factors as a product of polynomials of degree 1, we can repeat polynomials.

If all three polynomials are the same, we have 5 possibilities  $((x + a)^3, a \in \mathbb{F}_5)$ .

If one polynomial is repeated exactly two times, then we first choose which we will repeat; we can do this in 5 ways. Now we must choose one of the other 4; we can do this in 4 ways. Thus, there are  $5 \cdot 4 = 20$  such factorizations.

If neither polynomial is repeated, then we choose 3 polynomials from the set of 5 polynomials:  $\binom{5}{3} = 10$ .

Thus, there are  $5 + 20 + 10 = 35$  polynomials which we factor into linear polynomials.

On the other hand, there are  $10 \cdot 5 = 50$  monic polynomials of degree 1 which we factor as  $f(x) = p(x)q(x)$ , where  $p(x)$  is monic irreducible of degree 2, while  $q(x)$  is monic irreducible of degree 1.

So, there are  $35 + 50 = 85$  non-irreducible monic polynomials of degree 3. Since there are  $5^3 = 125$  monic polynomials of degree 3, we conclude that there are 40 monic irreducible polynomials of degree 3.

Now notice that if  $f(x)$  is irreducible, then  $af(x)$ ,  $a \in \mathbb{F}_5$ ,  $a \neq 0$ , is also irreducible. Similarly, if

$$f(x) = ax^3 + bx^2 + cx + d$$

is irreducible, then, if we write

$$f(x) = a(x^3 + a^{-1}bx^2 + a^{-1}cx + a^{-1}d) = ag(x),$$

$g(x)$  is monic irreducible polynomial.

Finally, we can now conclude that there are  $40 \cdot 4 = 160$  irreducible polynomials of degree 3 (in  $\mathbb{F}_5[x]$ , of course).

## Result

(a)

$$\begin{aligned} & x^3 + 2x + 1, \quad x^3 + 2x + 2, \quad x^3 + x^2 + 2, \quad x^3 + x^2 + x + 2, \\ & x^3 + x^2 + 2x + 1, \quad x^3 + 2x^2 + 1, \quad x^3 + 2x^2 + x + 1, \quad x^3 + 2x^2 + 2x + 2 \end{aligned}$$

(b)

$$\begin{aligned} & x^2 + x + 1, \quad x^2 + 4x + 1, \quad x^2 + 2, \quad x^2 + x + 2, \quad x^2 + 4x + 2, \\ & x^2 + 3, \quad x^2 + 2x + 3, \quad x^2 + 3x + 3, \quad x^2 + 2x + 4, \quad x^2 + 3x + 4 \end{aligned}$$

(c) 160.

## 13. a

(a)

We want it of degree  $n$  and to have all  $a_i$  except  $a_0$  as roots. A good idea is to set

$$p(x) = c(x - a_1) \cdots (x - a_n),$$

where  $c$  is some constant (notice that  $\deg p(x) = n$  and  $p(a_i) = 0$ ,  $i = 1, \dots, n$ ).

Now plug in  $a_0$ :

$$1 = p(a_0) = c(a_0 - a_1) \cdots (a_0 - a_n)$$

This makes sense since all  $a_i$  are distinct, so  $a_0 - a_i \neq 0$ ,  $i = 1, \dots, n$ . Finally,

$$c = \frac{1}{(a_0 - a_1) \cdots (a_0 - a_n)}$$

Therefore,

$$p(x) = \frac{(x - a_1) \cdots (x - a_n)}{(a_0 - a_1) \cdots (a_0 - a_n)}$$

This can also be written as

$$p(x) = \prod_{i=1}^n \frac{x - a_i}{a_0 - a_i}$$

(b)

Let  $p_i(x)$  be a polynomial of degree  $d$  such that  $p_i(a_i) = 1, p_i(a_j) = 0, j \neq i$  (such exists by (a)). Define

$$g(x) = b_0 p_0(x) + b_1 p_1(x) + \dots + b_d p_d(x)$$

Clearly  $g(a_i) = b_i, i = 0, 1, \dots, d$ . Also, since it is a linear combination of polynomials of degree  $d$ , its degree is  $\leq d$ .

Also, if we want to write it in terms of  $a_i$  and  $b_i$ , we can use the notation as in (a):

$$g(x) = \sum_{i=1}^d \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

Now suppose that  $h(x)$  is some other polynomial of degree  $\leq d$  such that  $h(a_i) = b_i, i = 0, 1, \dots, d$ . Define

$$f(x) = g(x) - h(x)$$

Notice that  $f(x)$  is of degree  $\leq d$ , and

$$f(a_i) = g(a_i) - h(a_i) = b_i - b_i = 0, \quad i = 0, 1, \dots, d$$

So,  $f(x)$  has  $d + 1$  roots. Since a nonzero polynomial of degree  $\leq d$  can have at most  $d$  roots (to be precise, the number of its roots cannot exceed the degree of said polynomial), we conclude that  $f(x) = 0$  (that is, that it is a zero polynomial). Finally,

$$g(x) - h(x) = 0 \implies g(x) = h(x)$$

Therefore, uniqueness is also proved.

## Result

3 of

(a)

$$p(x) = \frac{(x - a_1) \cdots (x - a_n)}{(a_0 - a_1) \cdots (a_0 - a_n)}$$

This can also be written as

$$p(x) = \prod_{i=1}^n \frac{x - a_i}{a_0 - a_i}$$

(b)

$$g(x) = \sum_{i=1}^d \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

To prove that it is unique, suppose that  $h(x)$  is some other polynomial with these properties, and observe

$$f(x) = g(x) - h(x).$$

14. a



Let  $p(x, y) = x^2 + y^2 - 1$ . Suppose that it is not irreducible. Then there exists  $f(x, y) \in \mathbb{C}[x, y]$  which is not a unit, nor is it associated with  $p(x, y)$ , and  $f(x, y)$  divides  $p(x, y)$ . That is, there exists some  $g(x, y) \in \mathbb{C}[x, y]$  such that

$$p(x, y) = f(x, y)g(x, y)$$

If  $f(x, y)$  is constant, then it is a unit.

If  $g(x, y)$  is constant, then it is a unit, so  $p(x, y)$  and  $f(x, y)$  are associated.

Thus,  $f(x, y)$  and  $g(x, y)$  are not constant. Since  $p(x, y)$  is of degree 2,  $f(x, y)$  and  $g(x, y)$  must be of degree 1. Thus,

$$f(x, y) = a_1x + b_1y + c_1$$

$$g(x, y) = a_2x + b_2y + c_2$$

From factorization  $p(x, y) = f(x, y)g(x, y)$ , we now get that

$$\{p = 0\} = \{f = 0\} \cup \{g = 0\}$$

However,  $\{p = 0\}$  is a circle, while  $\{f = 0\}$  and  $\{g = 0\}$  are lines. Thus, we get that a circle is a union of two lines, which is absurd.

Therefore,  $p(x, y) = f(x, y)g(x, y)$  cannot hold.

To conclude, we now get that  $p(x, y)$  is irreducible in  $\mathbb{C}[x, y]$ .

## Result

2 of 2

Suppose that it is not irreducible. From this it follows that a circle is a union of two lines (why?). Argue that this leads to a contradiction.

## 15. a

Consider the provided statement to explain what happens to given condition with reference to the Eisenstein criterion.

(a)

As provided condition is,  $\overline{f}$  is constant.

**Claim1:-**

Let  $f(x) = a_nx^n + \dots + a_0 \in \mathbb{Z}[x]$  and it is assumed that  $p$  is prime and  $p \in \mathbb{Z}$ . Then  $f$  is irreducible in  $\mathbb{Q}[x]$  if the condition which is provided below must satisfied,

$$(i) \quad p \nmid a_0$$

$$(ii) \quad p \mid a_i$$

$$(iii) \quad p^2 \nmid a_n$$

Where the value of  $i$  lies between the interval  $1 \leq i \leq n$  and condition (i) and (ii) hold if and only if  $\overline{f}$  is a nonzero constant.

**Proof:-**

If  $f = gh$  where the value of  $g, h$  is provided as below:

$$g = b_r x^r + \dots + b_0$$

$$h = c_s x^s + \dots + c_0$$

$$r + s = n$$

Since by (i) and (ii),

$$\begin{aligned}\overline{f} &= \overline{gh} \\ &= \overline{a_0} \\ &\neq 0\end{aligned}$$

As they must divide  $\overline{a_0}$  then it is seen that  $\overline{g} = \overline{b_0}$  and  $\overline{h} = \overline{c_0}$ .

Then  $p \mid b_r$  and  $p \mid c_s$ , when multiply both these equations so  $p^2 \mid b_r c_s = a_n$ .

(b)

Provided expression from textbook is,  $\overline{f} = x^n + \overline{b}x^{n-1}$

**Claim2:-**

Let  $f(x) = a_n x^n + \dots + a_0 \in \mathbb{Z}[x]$  and it is assumed that  $p$  is prime and  $p \in \mathbb{Z}$ . Then  $f$  is irreducible in  $\mathbb{Q}[x]$  if the condition which is provided below must satisfied,

$$(i) \ a_n \equiv 1 \pmod{p}$$

$$(ii) \ p \nmid a_{n-1}$$

$$(iii) \ p \nmid a_i$$

$$(iv) \ p \nmid a_0$$

Where the value of  $i$  lies between the interval  $0 \leq i \leq n-2$  and condition (i), (ii) and (iii) hold if and only if  $\overline{f} = x^n + \overline{b}x^{n-1}$ .

**Proof:-**

If  $f = gh$  where the value of  $g, h$  is provided as below:

$$g = b_r x^r + \dots + b_0$$

$$h = c_s x^s + \dots + c_0$$

$$r + s = n$$

Since by (i), (ii) and (iii),

$$\begin{aligned}\overline{f} &= \overline{gh} \\ &= \overline{a_{n-1}} x^{n-1} \\ &= x^{n-1} (x + \overline{a_{n-1}})\end{aligned}$$

The value of  $\overline{g}$  and  $\overline{h}$  is provided as below without loss of generality. As  $x + \overline{a_{n-1}}$  divides  $\overline{g}$  and  $\overline{g} = \overline{b_r} x^r (x + \overline{a_{n-1}})$ ,  $h = \overline{c_s} x^s$  therefore,  $p \mid c_0$  implies that  $p$  divides  $b_0 c_0$  which is equal to  $a_0$ .

16. a



We will prove that it is irreducible.

Suppose that

$$f(x) = g(x)h(x)$$

for some polynomials  $g(x), h(x) \in \mathbb{Q}[x]$ . We can multiply  $g(x)$  by the least common multiple of denominators of its coefficients, and divide it by the greatest common divisor of numerators of its coefficients, to get

$$f(x) = \tilde{g}(x)\tilde{h}(x),$$

where  $\tilde{g}(x) \in \mathbb{Z}[x]$  is primitive (notice that we must multiply  $h(x)$  with GCD and divide it by LCM to preserve the equality). Now we know that  $\tilde{g}(x)$  divides  $f(x)$  even in  $\mathbb{Z}[x]$ ; that is,

$$f(x) = \tilde{g}(x)q(x)$$

for some  $q(x) \in \mathbb{Z}[x]$ .

The residue is

$$\bar{f}(x) = x^{14} + 8x^{13} = x^{14} + 2x^{13}$$

and

$$\bar{f}(x) = \bar{g}(x)\bar{q}(x)$$

We have a factorization

$$x^{13}(x + 2)$$

Since  $\mathbb{F}_3[x]$  is a unique factorization domain, this factorization is unique. Thus, we must have

$$\bar{g}(x) = x^r$$

$$\bar{q}(x) = x^{13-r}(x + 2)$$

(it does not matter where we put  $(x + 2)$ ). Suppose that  $r \neq 0$  and  $r \neq 13$ . Then we would have that the constant terms of both  $\tilde{g}(x)$  and  $q(x)$  are multiples of 3 (since their residues have constant terms equal to zero). But this would mean that the constant term of their product  $f(x)$  is a multiple of 9, which is a contradiction since 3 is not a multiple of 9.

Thus, either  $r = 0$  or  $r = 13$ .

Suppose that  $r = 13$ . Then

$$\bar{g}(x) = x^{13}$$

$$\bar{q}(x) = x + 2$$

So, the degree of  $g(x)$  is at least 13, while the degree of  $q(x)$  is at least 1. Since the degree of  $f(x)$  is 14,  $g(x)$  must be of degree 13, while  $q(x)$  must be of degree 1. Thus,

$$q(x) = ax + b, \quad a \neq 0$$

This means that  $q(-b/a) = 0$ , so  $q$  has a rational root, and  $f(x)$  must now also have a rational root. By the rational root theorem, if  $x = m/n$  is a rational root of  $f(x)$ ,  $m$  must divide 3 (the constant term), while  $n$  must divide 1 (the leading coefficient). Thus, all candidates are  $\pm 3$ . However, neither 3 nor  $-3$  is a root of  $f(x)$ , meaning that  $f(x)$  has no rational roots.

So,  $r = 13$  is also impossible, so we must have  $r = 0$ . But this means that

$$\bar{g}(x) = 1, \quad \bar{q}(x) = x^{13}(x + 2)$$

As before, the degree of  $\tilde{g}(x)$  is at least 0, while the degree of  $q(x)$  is at least 14. Since the degree of their product  $f(x)$  is 14,  $\tilde{g}(x)$  is of degree 0, meaning that it is a unit in  $\mathbb{Q}[x]$ !

From this we conclude that if some polynomial  $p(x)$  divides  $f(x)$ , it is either a unit, or the quotient is a unit, meaning that  $p(x)$  and  $f(x)$  are associated. (Before, we proved that  $\tilde{g}(x)$  is a unit in  $\mathbb{Q}[x]$ . However, clearly  $g(x)$  is also a unit in  $\mathbb{Q}[x]$ , since  $g(x) = c\tilde{g}(x)$  for some  $c \in \mathbb{Q}$ !)

Finally, this means that  $f(x)$  is irreducible in  $\mathbb{Q}[x]$ !

## Result

It is irreducible in  $\mathbb{Q}[x]$ .

## 17. a

We will prove that it is irreducible.

Suppose that

$$f(x) = g(x)h(x)$$

for some polynomials  $g(x), h(x) \in \mathbb{Q}[x]$ . We can multiply  $g(x)$  by the least common multiple of denominators of its coefficients, and divide it by the greatest common divisor of numerators of its coefficients, to get

$$f(x) = \tilde{g}(x)\tilde{h}(x),$$

where  $\tilde{g}(x) \in \mathbb{Z}[x]$  is primitive (notice that we must multiply  $h(x)$  with GCD and divide it by LCM to preserve the equality). Now we know that  $\tilde{g}(x)$  divides  $f(x)$  even in  $\mathbb{Z}[x]$ ; that is,

$$f(x) = \tilde{g}(x)q(x)$$

for some  $q(x) \in \mathbb{Z}[x]$ .

The residue is

$$\bar{f}(x) = x^4 + 2x^3 + 3x^2 + 1$$

and

$$\bar{f}(x) = \bar{g}(x)\bar{q}(x)$$

Suppose that neither  $\bar{g}(x)$  nor  $\bar{q}(x)$  are constant. Furthermore, if for example  $\bar{g}(x)$  is of degree 1, that is,  $\bar{g}(x) = ax + b$ , for some  $a, b \in \mathbb{F}_4$ ,  $a \neq 0$ , then  $\bar{g}(-a^{-1}b) = 0$ , so  $\bar{g}(x)$  has a root in  $\mathbb{F}_4$ . But this means that  $\bar{f}(x)$  also has a root in  $\mathbb{F}_4$ , which we can directly check is not true (that is,  $\bar{f}(x)$  has no roots in  $\mathbb{F}_4$ ).

Since we arrive to the similar conclusion if we assume that  $\bar{q}(x)$  is linear, we know that the degrees of  $\bar{g}(x)$  and  $\bar{q}(x)$  are at least 2. However, their product is of degree 4, meaning that they must be of degree 2.

Let

$$\begin{aligned}\bar{g}(x) &= a_1x^2 + b_1x + c_1x \\ \bar{q}(x) &= a_2x^2 + b_2x + c_2x,\end{aligned}$$

where  $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{F}_4$ ,  $a_1, a_2 \neq 0$ . Furthermore, we can write

$$\bar{g}(x)\bar{q}(x) = (x^2 + e_1x + f_1)(d_2x^2 + e_2x + f_2),$$

where  $e_1 = a_1^{-1}b_1$ ,  $f_1 = a_1^{-1}c_1$ ,  $d_2 = a_1a_2$ ,  $e_2 = a_1b_2$ ,  $f_2 = a_1c_2$  (so, factor  $a_1$  out of the first polynomial, and multiply the second polynomial by it). Moreover,

$$(x^2 + e_1x + f_1)(d_2x^2 + e_2x + f_2) = d_2x^4 + (e_2 + e_1d_2)x^3 + (f_2 + e_1e_2 + f_1f_2)x^2 + (e_1f_2 + f_1e_2)x + f_1f_2$$

Now we have the equality of polynomials

$$x^4 + 2x^3 + 3x^2 + 1 = d_2x^4 + (e_2 + e_1d_2)x^3 + (f_2 + e_1e_2 + f_1f_2)x^2 + (e_1f_2 + f_1e_2)x + f_1f_2,$$

which yields the system of equations

$$\begin{aligned}d_2 &= 1 \\ e_2 + e_1d_2 &= 2 \\ f_2 + e_1e_2 + f_1f_2 &= 3 \\ e_1f_2 + f_1e_2 &= 0 \\ f_1f_2 &= 1\end{aligned}$$

So,  $d_2 = 1$  is clear. Plugging this into the second equation we get

$$e_1 + e_2 = 2$$

Now we manually solve  $f_1f_2 = 1$ :

$f_1 \cdot f_2$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Thus, there are two solutions:

$$\boxed{f_1 = f_2 = 1} \quad \text{or} \quad \boxed{f_1 = f_2 = 3}$$

Suppose that  $f_1 = f_2 = 1$ . Then

$$e_1 f_2 + e_2 f_1 = 0 \implies e_1 + e_2 = 0$$

But this contradicts the equation  $e_1 + e_2 = 2$ !

Similarly, if  $f_1 = f_2 = 3$ , then

$$e_1 f_2 + e_2 f_1 = 0 \implies 3(e_1 + e_2) = 0 \implies e_1 + e_2 = 0$$

Therefore, factorization

$$\bar{f}(x) = \bar{g}(x)\bar{q}(x)$$

with both  $\bar{g}(x)$  and  $\bar{q}(x)$  nonconstant does not exist. Without loss of generality, assume that  $\bar{g}(x)$  is of degree 0. Then  $\bar{q}(x)$  is of degree 4 (the degree of  $\bar{f}(x)$ ). But this means that the degree of  $\tilde{g}(x)$  is at least 0, while the degree of  $q(x)$  is at least 4. Since their product is  $f(x)$ , which is of degree 4, we conclude that  $q(x)$  is of degree 4, while  $\tilde{g}(x)$  is of degree 0. But this means that  $\tilde{g}(x)$  is a unit in  $\mathbb{Q}[x]$ !

Finally, this means that if some polynomial  $p(x)$  divides  $f(x)$ , then it is either a unit, or the quotient is a unit, making  $p(x)$  to be associated with  $f(x)$ . (Before, we proved that  $\tilde{g}(x)$  is a unit in  $\mathbb{Q}[x]$ . However, clearly  $g(x)$  is also a unit in  $\mathbb{Q}[x]$ , since  $g(x) = c\tilde{g}(x)$  for some  $c \in \mathbb{Q}$ !)

Therefore,  $f(x)$  is irreducible over  $\mathbb{Q}[x]$ .

## Result

It is irreducible.

18. a

Consider the provided statement to prove the given cyclotomic polynomial is irreducible.

Provided cyclotomic polynomial is as below and  $q = p^e, r = p^{e-1}$  with prime  $p$ .

$$\frac{(x^q - 1)}{(x^r - 1)}$$

[Comment](#)

Step 2 of 3 ^

As it is provided that  $p$  be prime and assume  $e \geq 1$ . Now it is to be shown that  $\phi_{p^e}$  is irreducible in  $\mathbb{Z}[x]$ . Therefore for making it irreducible in  $\mathbb{Q}[x]$  Gauss's Lemma is recalled, then

$$\begin{aligned} \phi_{p^e}(x) &= \phi_p(x^{p^{e-1}}) \\ &= \frac{x^{p^e} - 1}{x^{p^{e-1}} - 1} \end{aligned}$$

Therefore, given statement is in  $\mathbb{Z}[x]$ .

$$(x^{p^{e-1}} - 1)\phi_{p^e}(x) = x^{p^e} - 1$$

When the coefficients of the equation is reduced in modulo  $p$  then it is in the  $\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)[x]$ ,

$$(x^{p^{p-1}} - 1)\phi_{p^p}(x) = x^{p^p} - 1$$

Therefore, by lemma the obtained result is in  $\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)[x]$

$$(x-1)^{p^{p-1}}\phi_{p^p}(x) = (x-1)^{p^p}$$

Then,

$$\phi_{p^p}(x) = (x-1)^{p^{p-1}(p-1)}$$

It can be also say that, the obtained result is in  $\frac{\mathbb{Z}}{p^2\mathbb{Z}}$

$$\begin{aligned}\phi_{p^p}(1) &= \phi_p(1^p) \\ &= \phi_p(1) \\ &= p \neq 0\end{aligned}$$

Hence, from the Scho'nemann's Criterion  $\phi_{p^p}(x)$  is irreducible in  $\mathbb{Z}[x]$ . Therefore, it is irreducible in  $\mathbb{Q}[x]$  by Gauss's Lemma is **proved**.

19. a

(a)

First of all, this polynomial in  $\mathbb{F}_2$  is

$$p(x) = x^5 + x^4 + x^2 + 1$$

Notice that 1 is a root of this polynomial, so  $x - 1 = x + 1$  divides it. Thus, we need to find  $q(x) \in \mathbb{F}_2[x]$  such that

$$p(x) = (x + 1)q(x)$$

Since  $p(x)$  is of degree 5,  $q(x)$  is of degree 4. Thus,

$$q(x) = ax^4 + bx^3 + cx^2 + dx + e,$$

for some  $a, b, c, d, e \in \mathbb{F}_2$ . Furthermore,

$$(x + 1)q(x) = ax^5 + (a + b)x^4 + (b + c)x^3 + (c + d)x^2 + (d + e)x + e$$

Thus,  $p(x) = (x + 1)q(x)$  if and only if

$$\begin{aligned}a &= 1 \\ a + b &= 1 \\ b + c &= 0 \\ c + d &= 1 \\ d + e &= 0 \\ e &= 1\end{aligned}$$

From this system, we get the solution

$$\boxed{a = 1, b = 0, c = 0, d = 1, e = 1}$$

Therefore,

$$p(x) = (x + 1)(x^4 + x + 1)$$



Now suppose that we can factor  $x^4 + x + 1$  furthermore:

$$x^4 + x + 1 = q_1(x)q_2(x),$$

where  $q_1(x)$  and  $q_2(x)$  are not constant. If  $q_1(x)$  is of degree 1, then  $q_1(x) = ax + b$ , and  $q_1(-a^{-1}b) = 0$ . Thus,  $q_1(x)$  has a root, so  $x^4 + x + 1$  must also have a root in  $\mathbb{F}_2$ . But we see that it has no roots in  $\mathbb{F}_2$ !

We arrive to the same conclusion if we assume that  $q_2(x)$  is of degree 1.

Thus,  $q_1(x)$  and  $q_2(x)$  must be of degree 2 or more. Since their product is of degree 4, they must be of degree 2. Furthermore, they must be irreducible (otherwise, if  $q_1(x)$  was not irreducible, it could be written as a product of polynomials of degree 1, so it would have a root, and  $x^4 + x + 1$  would also have a root; contradiction).

The only irreducible element of degree 2 in  $\mathbb{F}_2[x]$  is  $x^2 + x + 1$  (see (12.4.4) on page 373). So, we must have that

$$x^4 + x + 1 = (x^2 + x + 1)^2$$

However,

$$(x^2 + x + 1)^2 = x^4 + x^2 + 1 \neq x^4 + x + 1$$

Therefore,  $x^4 + x + 1$  cannot be factored further (it is irreducible). Thus,

$$p(x) = (x + 1)(x^4 + x + 1)$$

**(b)**

Denote this polynomial as  $p(x)$ . Then notice that  $p(-5) = 0$  (in  $\mathbb{Z}/16\mathbb{Z}$ ). Therefore,  $x + 5$  divides  $p(x)$ , so we need to find  $q(x) \in \mathbb{Z}/16\mathbb{Z}[x]$  such that

$$p(x) = (x + 5)q(x)$$

Since  $p(x)$  is of degree 5,  $q(x)$  is of degree 4. Thus,

$$q(x) = ax^4 + bx^3 + cx^2 + dx + e,$$

for some  $a, b, c, d, e \in \mathbb{Z}/16\mathbb{Z}$ . Furthermore,

$$(x + 5)q(x) = ax^5 + (5a + b)x^4 + (5b + c)x^3 + (5c + d)x^2 + (5d + e)x + 5e$$

Thus,  $p(x) = (x + 5)q(x)$  if and only if

$$\begin{aligned} a &= 1 \\ 5a + b &= -1 \\ 5b + c &= 0 \\ 5c + d &= -1 \\ 5d + e &= 0 \\ 5e &= -1 \end{aligned}$$

From this system, we get the solution

$$a = 1, b = 10, c = 14, d = 9, e = 3$$

Therefore,

$$p(x) = (x + 5)(x^4 + 10x^3 + 14x^2 + 9x + 3)$$



Now suppose that we can factor  $x^4 + 10x^3 + 14x^2 + 9x + 3$  further; that is, that there exist polynomials  $f(x), g(x) \in \mathbb{Z}/16\mathbb{Z}$  such that

$$x^4 + 10x^3 + 14x^2 + 9x + 3 = f(x)g(x)$$

But the residue of  $x^4 + 10x^3 + 14x^2 + 9x + 3$  in  $\mathbb{F}_2[x]$  is

$$x^4 + x + 1,$$

so we would now have that

$$x^4 + x + 1 = \bar{f}(x)\bar{g}(x)$$

Thus, the factorization of  $x^4 + 10x^3 + 14x^2 + 9x + 3$  over  $\mathbb{Z}/16\mathbb{Z}$  would induce a factorization of  $x^4 + x + 1$  in  $\mathbb{F}_2$ . But this is a contradiction! In **(a)** we concluded that we cannot factorize it further.

Therefore,

$$p(x) = (x + 5)(x^4 + 10x^3 + 14x^2 + 9x + 3)$$

**(c)**

Similarly to **(b)**, if we can factor this polynomial in a way for one of the factors to be of degree 2 or 3, it would induce a factorization over  $\mathbb{F}_2[x]$  where some factor of  $x^4 + x + 1$  would be of degree 2 or 3. Thus, the only possible factorization is of the form

$$p(x) = (ax + b)q(x),$$

with  $q(x)$  of degree 4. But this means that  $p(x)$  has a rational root!

By the rational root theorem, if  $m/n$  is a rational root of  $p(x)$ , then  $m$  divides  $-1$  (the constant term), and  $n$  divides 1 (the leading coefficient). Thus, all possible rational roots are  $\pm 1$ . However, neither 1 nor  $-1$  is a root of  $p(x)$ , hence a contradiction.

Therefore, we conclude that  $p(x)$  is irreducible over  $\mathbb{Q}$ .

**Result**

6 of 6

$$\textbf{(a)} \ (x + 1)(x^4 + x + 1)$$

$$\textbf{(b)} \ (x + 5)(x^4 + 10x^3 + 14x^2 + 9x + 3)$$

**(c)** This polynomial is irreducible over  $\mathbb{Q}$ .

## Section 5

1. a

For this exercise, we first define a mapping

$$N : \mathbb{Z}[i] \rightarrow \mathbb{N} \cup \{0\}, \quad N(a + bi) = a^2 + b^2$$

It is easy to see that

$$\begin{aligned} N((a + bi)(c + di)) &= N((ac - bd) + (ad + bc)i) \\ &= (ac - bd)^2 + (ad + bc)^2 \\ &= a^2c^2 + a^2d^2 + b^2d^2 + b^2c^2 \end{aligned}$$

and

$$N(a + bi)N(c + di) = (a^2 + b^2)(c^2 + d^2) = a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2$$

Therefore,

$$N((a + bi)(c + di)) = N(a + bi)N(c + di)$$

This gives us a powerful tool; if  $y$  divides  $x$ , then

$$x = yz \implies N(x) = N(y)N(z),$$

so  $N(y)$  divides  $N(x)$ . Furthermore,  $N(x) = 1$  if and only if  $x = \pm 1, \pm i$ , which are of no interest to us (they are units), so we look for divisors of  $x$  with  $N(y) > 1$ .

Furthermore,  $a + bi$ , with  $a \neq 0$  and  $b \neq 0$ , is prime if  $N(a + bi) = a^2 + b^2$  is a prime integer (this is a variation of **Theorem 12.5.2 (b)**).

**(a)**

First of all,

$$N(1 - 3i) = 10,$$

so we look for  $a + bi \in \mathbb{Z}[i]$  such that  $N(a + bi) = 2$  or  $N(a + bi) = 5$ . Notice that there are only 4 numbers such that  $N(a + bi) = 2$ :

$$1 + i, 1 - i, -1 + i, -1 - i$$

Moreover,

$$(1 + i)(-1 - 2i) = 1 - 3i$$

Since  $N(1 + i) = 2$ , which is prime, and  $N(-1 - 2i) = 5$ , which is also prime, we conclude that this is a prime factorization:

$$\boxed{1 - 3i = (1 + i)(-1 - 2i)}$$

(b)

First of all,

$$N(10) = 100,$$

so we look for  $a + bi \in \mathbb{Z}[i]$  such that  $N(a + bi) = 2$  or  $N(a + bi) = 5$  (there is no sense in finding  $a + bi$  such that, for example,  $N(a + bi) = 10$ , since we only look for prime elements, and no prime integer  $p$ ,  $p \equiv 3$  modulo 4 divides 10, so all prime divisors of 10 will be of the form  $a + bi$ ,  $a \neq 0$ ,  $b \neq 0$ ). Notice that there are only 4 numbers such that  $N(a + bi) = 2$ :

$$1 + i, 1 - i, -1 + i, -1 - i$$

Moreover,

$$(1 + i)(5 - 5i) = 10$$

Furthermore,

$$5 - 5i = (1 - i) \cdot 5$$

So, we only need to factor 5 further (because  $1 + i$  and  $1 - i$  are prime). However, notice that

$$(2 - i)(2 + i) = 5,$$

and  $N(2 - i) = N(2 + i) = 5$ , so they are also prime. Finally, we get the following prime factorization:

$$10 = (1 + i)(1 - i)(2 + i)(2 - i)$$

(c)

We can factor this faster:

$$6 + 9i = 3(2 + 3i)$$

3 is prime in  $\mathbb{Z}[i]$  (**Theorem 12.5.2 (c)**), and  $N(2 + 3i) = 13$ , which is prime, so it is also a prime in  $\mathbb{Z}[i]$ . Thus, the required prime factorization is

$$6 + 9i = 3(2 + 3i)$$

(d)

First of all,

$$N(7+i) = 50,$$

so we look for  $a+bi \in \mathbb{Z}[i]$  such that  $N(a+bi) = 2$  or  $N(a+bi) = 5$  (there is no sense in finding  $a+bi$  such that, for example,  $N(a+bi) = 10$ , since we only look for prime divisors, and no integer  $d$  other than 1 and  $-1$  divide  $7+i$  since  $d$  must divide both 7 and 1 by **Lemma 12.5.1**, so all prime divisors will be of the form  $a+bi$  with  $a \neq 0$  and  $b \neq 0$ ). Notice that there are only 4 numbers such that  $N(a+bi) = 2$ :

$$1+i, 1-i, -1+i, -1-i$$

Moreover,

$$(1+i)(4-3i) = 7+i$$

Since  $N(1+i) = 2$ , it is prime. Furthermore,  $N(4-3i) = 25$ , so we need to find  $a+bi$  such that  $N(a+bi) = a^2+b^2 = 5$ . Clearly the only candidates are

$$2+i, 2-i, -2+i, -2-i, 1+2i, 1-2i, -1+2i, -1-2i$$

Now notice that

$$(2+i)(1-2i) = 4-3i$$

Finally, the required prime factorization is

$$7+i = (1+i)(2+i)(1-2i)$$

## Result

$$(a) (1+i)(-1-2i)$$

$$(b) (1+i)(1-i)(2+i)(2-i)$$

$$(c) 3(2+3i)$$

$$(d) (1+i)(2+i)(1-2i)$$

## 2. a

The general strategy is similar as with the Euclidean Algorithm. We divide  $a+bi$  by  $c+di$  in the following way:

$$\frac{a+bi}{c+di} \cdot \frac{c-di}{c+di} = k+li,$$

with  $k, l \in \mathbb{Q}$ . Now for the quotient we pick  $q = m+ni$ , with  $m, n \in \mathbb{Z}$  such that  $|m-k| \leq 1/2$  and  $|n-l| \leq 1/2$ .

Of course, for the remainder we have  $r = (a+bi) - q(c+di)$ .

The proof that

$$\gcd(a+bi, c+di) = \gcd(c+di, r)$$

is the same as with integers.

(a)

$$\frac{11+7i}{4+7i} \cdot \frac{4-7i}{4-7i} = \frac{93-49i}{65} = \frac{93}{65} - \frac{49}{65}i$$

Therefore, we set  $q_1 = 1 - i$ , and

$$r_1 = 11 + 7i - (1 - i)(4 + 7i) = 11 + 7i - (11 + 3i) = 4i$$

Now,

$$\frac{4+7i}{r_1} = \frac{4+7i}{4i} \cdot \frac{-4i}{-4i} = \frac{28-16i}{16} = \frac{7}{4} - i$$

Therefore, we set  $q_2 = 2 - i$ , and

$$r_2 = 4 + 7i - 4i(2 - i) = 4 + 7i - 4 - 8i = -i$$

Finally,

$$\frac{r_1}{r_2} = \frac{4i}{-i} = -4,$$

so  $q_3 = -4$ ,  $r_3 = 0$ , meaning that

$$\boxed{\gcd(11+7i, 4+7i) = -i}$$

Moreover, we can prove that 1 is also gcd (because  $-i$  is a unit). Truly, 1 divides  $11 + 7i$  and  $4 + 7i$ . Suppose that  $d$  divides  $11 + 7i$  and  $4 + 7i$ . Since  $-i$  is a gcd,  $d$  divides  $-i$ , so

$$-i = de$$

for some  $e \in \mathbb{Z}[i]$ . Multiplying the equality by  $i$  we get

$$1 = d(ei),$$

so  $d$  divides 1, making 1 the gcd of  $11 + 7i$  and  $4 + 7i$ .

(b)

$$\frac{11+7i}{8+i} \cdot \frac{8-i}{8-i} = \frac{95+45i}{65} = \frac{95}{65} + \frac{45}{65}i$$

Therefore, we set  $q_1 = 1 + i$ , and

$$r_1 = 11 + 7i - (8 + i)(1 + i) = 11 + 7i - (7 + 9i) = 4 - 2i$$

Now,

$$\frac{8+i}{r_1} = \frac{8+i}{4-2i} \cdot \frac{4+2i}{4+2i} = \frac{30+20i}{20} = \frac{3}{2} + i$$

Therefore, we set  $q_2 = 1 + i$  (or  $2 + i$ ), and

$$r_2 = 8 + i - (1 + i)(4 - 2i) = 8 + i - (6 + 2i) = 2 - i$$

Finally,

$$\frac{r_1}{r_2} = \frac{4-2i}{2-i} = 2,$$

so  $q_3 = 2$ ,  $r_3 = 0$ , meaning that

$$\boxed{\gcd(11+7i, 8+i) = 2-i}$$

(c)

$$\frac{18-i}{3+4i} \cdot \frac{3-4i}{3-4i} = \frac{50-75i}{25} = 2-3i$$

Therefore, we set  $q_1 = 2 - 3i$ , and

$$r_1 = 0$$

Therefore, here the situation is trivial:

$$\gcd(18-i, 3+4i) = 2-3i$$

## Result

(a)  $-i$  (or 1)

(b)  $2 - i$

(c)  $2 - 3i$

## 3. a

Since  $\mathbb{Z}[i]$  is a principal ideal domain,

$$(3+4i, 4+7i) = (d),$$

for some  $d \in \mathbb{Z}[i]$ . Now we will prove that  $d = \gcd(3+4i, 4+7i)$ .

First of all,  $3+4i \in (d)$ , so  $3+4i = de$  for some  $e \in \mathbb{Z}[i]$ , so  $d$  divides  $3+4i$ . Arguing similarly we get that  $d$  divides  $4+7i$ .

Now suppose that  $c$  is a common divisor of  $3+4i$  and  $4+7i$ . Since  $d \in (3+4i, 4+7i)$ ,  $d = a(3+4i) + b(4+7i)$  for some  $a, b \in \mathbb{Z}[i]$ . But now it becomes clear that  $c$  divides  $d$ .

Thus,  $d = \gcd(3+4i, 4+7i)$ .

Now we need to find the greatest common divisor. The general strategy is similar as with the Euclidean Algorithm. We divide  $a+bi$  by  $c+di$  in the following way:

$$\frac{a+bi}{c+di} \cdot \frac{c-di}{c+di} = k+li,$$

with  $k, l \in \mathbb{Q}$ . Now for the quotient we pick  $q = m+ni$ , with  $m, n \in \mathbb{Z}$  such that  $|m-k| \leq 1/2$  and  $|n-l| \leq 1/2$ .

Of course, for the remainder we have  $r = (a+bi) - q(c+di)$ .

The proof that

$$\gcd(a+bi, c+di) = \gcd(c+di, r)$$

is the same as with integers.



$$\frac{4+7i}{3+4i} \cdot \frac{3-4i}{3-4i} = \frac{40+5i}{25} = \frac{8}{5} + \frac{1}{5}i$$

Therefore, we set  $q_1 = 2$ , and

$$r_1 = 4 + 7i - (3 + 4i)(2) = 4 + 7i - (6 + 8i) = -2 - i$$

Now,

$$\frac{3+4i}{r_1} = \frac{3+4i}{-2-i} \cdot \frac{-2+i}{-2+i} = \frac{-10-5i}{5} = -2-i$$

Therefore, we set  $q_2 = -2 - i$ , and

$$r_2 = 3 + 4i - (-2 - i)^2 = 3 + 4i - (3 + 4i) = 0$$

Therefore,

$$\gcd(4 + 7i, 3 + 4i) = -2 - i$$

and

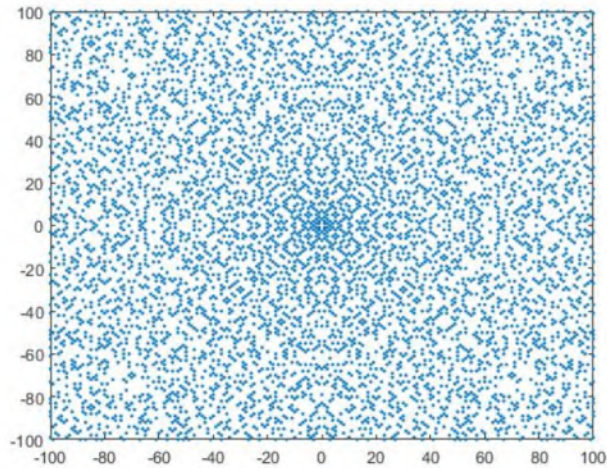
$$(4 + 7i, 3 + 4i) = (-2 - i)$$

## Result

$$-2 - i$$

4. a

From **Theorem 12.5.2** we conclude that  $\pi = a + bi$  is a Gaussian prime if and only if  $a \neq 0, b \neq 0, a^2 + b^2$  is a prime integer, or  $b = 0, a$  is a prime integer congruent to 3 modulo 4, or  $a = 0, b$  is a prime integer congruent to 3 modulo 4. From this, we can easily get the following picture:



## Result

2 of 2

Hint:  $\pi = a + bi$  is a Gaussian prime if and only if  $a^2 + b^2$  is a prime integer, or  $b = 0, a$  is a prime integer congruent to 3 modulo 4, or  $a = 0, b$  is a prime integer congruent to 3 modulo 4.

5. a

$\pi$  and  $\bar{\pi}$  are associates.

Assume that  $\pi$  and  $\bar{\pi}$  are associates. Let  $\pi = a + bi$ . Then  $\bar{\pi} = a - bi$ , and there exists some unit  $u \in \mathbb{Z}[i]$  such that

$$\pi = u\bar{\pi} \implies a + bi = u(a - bi)$$

Therefore,

$$a^2 + b^2 = (a + bi)(a - bi) = u\bar{\pi}^2 = u(a - bi)^2 = u(a^2 - b^2 - 2abi)$$

There are only 4 units in  $\mathbb{Z}[i]$ :  $\pm 1$  and  $\pm i$ . If  $u = \pm 1$ , then

$$a^2 + b^2 = \pm(a^2 - b^2 - 2abi)$$

Since  $a^2 + b^2 \in \mathbb{R}$ , we must have  $-2ab = 0$ ; so,  $a = 0$  or  $b = 0$ .

If  $a = 0$ , then  $\pi\bar{\pi} = \mp b^2$ . Furthermore,  $\pi\bar{\pi} > 0$ , so  $\pi\bar{\pi} = b^2$ . By **Theorem 12.5.2**, since  $b^2$  is clearly not a prime integer, we conclude that  $b$  is prime. But now  $\pi = bi$ , and  $i$  is a unit, so  $\pi$  is associated with  $b$ , hence  $\pi$  is associated with the integer prime!

If  $b = 0$ , then  $\pi\bar{\pi} = \pm a^2$ . Furthermore,  $\pi\bar{\pi} > 0$ , so  $\pi\bar{\pi} = a^2$ . By **Theorem 12.5.2**, since  $a^2$  is clearly not a prime integer, we conclude that  $a$  is prime. But now  $\pi = a$ , so  $\pi$  is a integer prime!

Now assume that  $u = i$ . Then

$$a^2 + b^2 = (a^2 - b^2)i + 2ab$$

Since  $a^2 + b^2 \in \mathbb{R}$ , we must have that  $a^2 - b^2 = 0$ , so  $a = \pm b$ . If  $a = -b$ , then the above equality yields

$$b^2 + b^2 = -2b^2,$$

which is clearly a contradiction. Thus,  $a = b$ . Now we have that

$$\pi\bar{\pi} = 2b^2$$

By **Theorem 12.5.2 (a)** we now conclude that  $2b^2$  must be either an integer prime, or a square of an integer prime. Notice that it cannot possibly be a square of an integer prime (nor a square of any integer!), so  $2b^2$  is a prime integer. But there is only one possibility:  $2b^2 = 2$ . Thus,

$$\pi\bar{\pi} = 2$$

Now assume that  $u = -i$ . Then

$$a^2 + b^2 = -(a^2 - b^2)i - 2ab$$

Since  $a^2 + b^2 \in \mathbb{R}$ , we must have that  $a^2 - b^2 = 0$ , so  $a = \pm b$ . If  $a = b$ , then the above equality yields

$$b^2 + b^2 = -2b^2,$$

which is clearly a contradiction. Thus,  $a = -b$ . Now we have that

$$\pi\bar{\pi} = 2b^2$$

By **Theorem 12.5.2 (a)** we now conclude that  $2b^2$  must be either an integer prime, or a square of an integer prime. Notice that it cannot possibly be a square of an integer prime (nor a square of any integer!), so  $2b^2$  is a prime integer. But there is only one possibility:  $2b^2 = 2$ . Thus,

$$\pi\bar{\pi} = 2$$

$\pi$  is an associate of an integer prime, or  $\pi\bar{\pi} = 2$

If  $\pi$  is an associate of an integer prime  $p$ , then there exists a unit  $u \in \mathbb{Z}[i]$  such that

$$\pi = up$$

As before, there are only 4 units in  $\mathbb{Z}[i]$ :  $u = \pm 1, \pm i$ . If  $u = \pm 1$ , then

$$\pi = \pm p \implies \bar{\pi} = \pm p = \pm \pi,$$

so  $\pi$  and  $\bar{\pi}$  are associated (since  $\pm 1$  is a unit).

If  $u = \pm i$ , then

$$\pi = \pm ip \implies \bar{\pi} = \mp ip = \mp \pi$$

Therefore,  $\pi$  and  $\bar{\pi}$  are associated.

Now suppose that  $\pi\bar{\pi} = 2$ . Let  $\pi = a + bi$ . Then

$$2 = \pi\bar{\pi} = (a + bi)(a - bi) = a^2 + b^2$$

Since  $a, b \in \mathbb{Z}$ ,  $a = \pm 1$  and  $b = \pm 1$ , hence we have only 4 possibilities:

$$\pi_1 = 1 + i, \pi_2 = 1 - i, \pi_3 = -1 + i, \pi_4 = -1 - i$$

If  $\pi = 1 + i$ , then

$$\bar{\pi} = 1 - i = -i(1 + i) = -i\pi,$$

so  $\pi$  and  $\bar{\pi}$  are associated since  $-i$  is a unit.

If  $\pi = 1 - i$ , then

$$\bar{\pi} = 1 + i = i(1 - i) = i\pi$$

Thus, they are associated, since  $i$  is a unit.

If  $\pi = -1 + i$ , then

$$\bar{\pi} = -1 - i = i(-1 + i) = i\pi$$

They are once again associated.

If  $\pi = -1 - i$ , then

$$\bar{\pi} = -1 + i = -i(-1 - i) = -i\pi$$

Thus, they are associated.

## Result

4 of 4

First part:  $\pi$  and  $\bar{\pi}$  are associated.

Write  $\pi = a + bi$ . Also,  $\pi = u\bar{\pi}$ , where  $u = \pm 1, \pm i$  (why?). If  $u = \pm 1$ , conclude that  $a = 0$  or  $b = 0$ . If  $u = \pm i$ , conclude that  $\pi\bar{\pi} = 2$  (**Theorem 12.5.2 (a)** can be useful).

Second part:  $\pi$  is associated with an integer prime.

Then  $\pi = up$ , where  $u = \pm 1, \pm i$  (why?). Now it is easy to show directly that  $\pi$  and  $\bar{\pi}$  are associated.

Final part. Assume that  $\pi\bar{\pi} = 2$ . Let  $\pi = a + bi$ . Then  $a^2 + b^2 = 2$ , so  $a = \pm 1, b = \pm 1$ . Show that now  $\pi$  and  $\bar{\pi}$  are associated (check each case manually).

We will first show that  $R = \mathbb{Z}[\sqrt{-3}] \approx \mathbb{Z}[x]/(x^2 + 3)$ . Observe the mapping  $\mathbb{Z}[x] \rightarrow R$  defined by  $x \mapsto \sqrt{-3}$ . It is clearly surjective, since for every  $a + b\sqrt{-3}$ ,  $a, b \in \mathbb{Z}$ , we have that  $f(x) = a + bx \in \mathbb{Z}[x]$ , and  $f(x) \mapsto a + b\sqrt{-3}$ .

Now we want to find its kernel  $K$ . Let  $f(x)$  be in the kernel. Then  $f(\sqrt{-3}) = 0$ . Hence,  $\sqrt{-3}$  is a root of  $f(x)$ . However,  $\sqrt{-3} = i\sqrt{3}$ , and since  $f(x)$  has real coefficients (it has integer coefficients, which are real), then by the complex conjugate theorem we conclude that we also have that  $-i\sqrt{3}$  is a root of  $f(x)$ .

Divide  $f(x)$  by  $x^2 + 3$  (we can do that in  $\mathbb{Z}[x]$  since the leading coefficient of  $x^2 + 3$  is 1, which is a unit in  $\mathbb{Z}$ ):

$$f(x) = (x^2 + 3)q(x) + r(x),$$

where  $q(x), r(x) \in \mathbb{Z}[x]$ , and  $r(x)$  is of degree 1 or less.

Now, if we denote  $g(x) = x^2 + 3$ , then clearly  $g(\sqrt{-3}) = 0$  and  $g(-i\sqrt{-3}) = 0$ . Therefore,

$$0 = f(\sqrt{-3}) = q(\sqrt{-3})g(\sqrt{-3}) + r(\sqrt{-3}) = r(\sqrt{-3})$$

$$0 = f(-i\sqrt{3}) = q(-i\sqrt{3})g(-i\sqrt{3}) + r(-i\sqrt{3}) = r(-i\sqrt{3})$$

Therefore,  $r(x)$  is a polynomial with complex coefficients of degree 1 or less which has two different roots. But this means that  $r(x)$  is a zero polynomial:  $r(x) = 0$ ! Therefore,

$$f(x) = (x^2 + 3)q(x)$$

So,  $f(x) \in (x^2 + 3)$ , so

$$K \subseteq (x^2 + 3)$$

For the other inclusion, let  $f(x) \in (x^2 + 3)$ . Then  $f(x) = g(x)(x^2 + 3)$ , for some  $g(x) \in \mathbb{Z}[x]$ . But now  $f(\sqrt{-3}) = 0$  is clear, so  $f(x) \in K$ . Therefore,

$$K = (x^2 + 3)$$

By the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(x^2 + 3) \approx R$$

By the symmetry of the relation "to be isomorphic to", we conclude that

$$R \approx \mathbb{Z}[x]/(x^2 + 3)$$

Consider the ring  $\bar{R} = \mathbb{Z}[\sqrt{3}]/(p)$ . The point is that we can get from  $\mathbb{Z}[x]$  to  $\bar{R}$  on two "paths" using the following diagram:

$$\begin{array}{ccc} \mathbb{Z}[x] & \xrightarrow[\text{kill } p]{\text{kill } x^2+3} & \mathbb{F}_p[x] \\ \downarrow \text{kill } x^2+3 & & \downarrow \text{kill } x^2+3 \\ R & \xrightarrow[\text{kill } p]{\text{kill } x^2+3} & \bar{R} \end{array}$$



From **Corollary 12.2.9 (b)**,  $p$  is prime in  $R$  if and only if it is irreducible in  $R$ . Truly,  $R$  is an Euclidean domain with  $\sigma(a + bi) = a^2 + 3b^2$ , so by **Proposition 12.2.7** it is a principal ideal domain. By **Corollary 12.2.9 (c)**,  $p$  is irreducible in  $R$  if and only if  $(p)$  is a maximal ideal in  $R$ . By **Proposition 11.8.2 (b)** this holds if and only if  $\overline{R}$  is a field.

On the other hand, by **Proposition 11.8.2 (b)** this holds if and only if  $(x^2 + 3)$  is a maximal ideal in  $\mathbb{F}_p[x]$ . Since  $\mathbb{F}_p[x]$  is a principal ideal domain, by **Corollary 12.2.9 (c)** we conclude that this is true if and only if  $x^2 + 3$  is irreducible in  $\mathbb{F}_p[x]$ .

Thus, putting everything together, we have proven precisely that  $p$  is prime in  $R$  if and only if  $x^2 + 3$  is irreducible in  $\mathbb{F}_p[x]$ , as required.

## Result

4 of 4

Hint: observe  $\overline{R} = R/(p)$ . Notice that we can get from  $\mathbb{Z}[x]$  to  $\overline{R}$  in two ways: first killing  $p$ , then killing  $x^2 + 3$ , or first killing  $x^2 + 3$ , then killing  $p$ .

## 7. a

If  $p$  is congruent to 3 modulo 4, then it is a Gauss prime (**Theorem 12.5.2 (c)**), so  $\mathbb{Z}[i]/(p)$  is a field by **Lemma 12.5.3 (b)**. Also,  $\mathbb{F}_p[x]/(x^2 + 1)$  is isomorphic to  $\mathbb{Z}[i]/(p)$  by diagram (12.5.4), so it is also a field. Moreover,  $\mathbb{F}_p[x]/(x^2 + 1)$  consists of elements of the form  $ax + b + (x^2 + 1)$ ,  $a, b \in \mathbb{F}_p$  (truly, when we have  $f(x) + (x^2 + 1)$ , we divide  $f(x)$  by  $x^2 + 1$  to get a linear polynomial  $r(x)$  such that  $f(x) + (x^2 + 1) = q(x)(x^2 + 1) + r(x) + (x^2 + 1) = r(x) + (x^2 + 1)$ , since  $q(x)(x^2 + 1) \in (x^2 + 1)$ ). This means that  $\mathbb{F}_p[x]/(x^2 + 1)$  is a field with  $p^2$  elements, so  $\mathbb{F}_p[x]/(x^2 + 1) \approx \mathbb{F}_{p^2} \implies \mathbb{Z}[i]/(p) \approx \mathbb{F}_{p^2}$ .

If  $p$  is congruent to 1 modulo 4, or  $p = 2$ , then it is not a Gauss prime (**Theorem 12.5.2 (c)**), so by **Lemma 12.5.3 (c)**  $x^2 + 1$  is not irreducible in  $\mathbb{F}_p[x]$ . Thus,  $x^2 + 1 = (x + a)(x + b)$ , for some  $a, b \in \mathbb{F}_p$ . Moreover,  $a \neq 0$  and  $b \neq 0$ , since 0 is not a root of  $x^2 + 1$ . Furthermore, from the diagram (12.5.4) we conclude that  $\mathbb{Z}[i]/(p) \approx \overline{R} \approx \mathbb{F}_p[x]/(x^2 + 1)$ , so  $\mathbb{Z}[i]/(p) \approx \mathbb{F}_p[x]/(x^2 + 1)$ .

If  $p = 2$ , we cannot simplify this further. So we assume that  $p \neq 2$ . Therefore,  $a \neq b$ . Truly, if  $a = b$ , then  $x^2 + 1 = (x + a)^2 = x^2 + 2ax + a^2$ . But this means that  $2a = 0$  in  $\mathbb{F}_p$ , so  $2a \equiv 0$  modulo  $p$ . But this is absurd since  $2a < 2p$ , so we would have that  $a = 0$  (which is not since  $x^2 + 1$  does not have 0 as a root) or  $2a - p = 0 \implies p = 2a$ , which is absurd since  $p$  is an odd prime.

Now we will prove that  $\mathbb{F}_p[x]/(x^2 + 1) \approx \mathbb{F}_p[x]/(x + a) \times \mathbb{F}_p[x]/(x + b)$ . Define a mapping

$$\varphi : \mathbb{F}_p[x] \rightarrow \mathbb{F}_p[x]/(x + a) \times \mathbb{F}_p[x]/(x + b)$$

by

$$\varphi(f(x)) = (f(x) + (x + a), f(x) + (x + b))$$

It is a homomorphism:

$$\begin{aligned}
\varphi(f(x) + g(x)) &= ((f(x) + g(x)) + (x + a), (f(x) + g(x)) + (x + b)) \\
&= ((f(x) + (x + a)) + (g(x) + (x + a)), (f(x) + (x + b)) + (g(x) + (x + b))) \\
&= (f(x) + (x + a), f(x) + (x + b)) + (g(x) + (x + a), g(x) + (x + b)) \\
&= \varphi(f(x)) + \varphi(g(x))
\end{aligned}$$

$$\begin{aligned}
\varphi(f(x)g(x)) &= ((f(x)g(x)) + (x + a), (f(x)g(x)) + (x + b)) \\
&= ((f(x) + (x + a))(g(x) + (x + a)), (f(x) + (x + b))(g(x) + (x + b))) \\
&= (f(x) + (x + a), f(x) + (x + b))(g(x) + (x + a), g(x) + (x + b)) \\
&= \varphi(f(x))\varphi(g(x))
\end{aligned}$$

$$\varphi(1) = (1 + (x + a), 1 + (x + b))$$

It is also surjective. All elements of  $\mathbb{F}_p[x]/(x + a)$  are of the form  $c + (x + a)$ , with  $c \in \mathbb{F}_p$ . Truly, when we have  $f(x) + (x + a)$ , we can divide  $f(x)$  by  $x + a$  to get

$$f(x) + (x + a) = g(x)(x + a) + r(x) + (x + a) = r(x) + \underbrace{g(x)(x + a)}_{\in (x+a)} + (x + a) = r(x) + (x + a)$$

Also,  $r(x)$  is of degree 0, so  $r(x) = c$ , for some  $c \in \mathbb{F}_p$ .

Similarly, all elements of  $\mathbb{F}_p[x]/(x + b)$  are of the form  $d + (x + b)$ ,  $d \in \mathbb{F}_p$ .

Now observe

$$f(x) = d(a - b)^{-1}(x + a) + c(b - a)^{-1}(x + b)$$

(it is well-defined since  $a \neq b$ , so  $a - b$  and  $b - a$  have inverses in the field  $\mathbb{F}_p$ ).

Then, since  $d(a - b)^{-1}(x + a) \in (x + a)$  and  $c(b - a)^{-1}(x + b) \in (x + b)$ ,

$$\varphi(f(x)) = (c(b - a)^{-1}(x + b) + (x + a), d(a - b)^{-1}(x + a) + (x + b))$$

Furthermore,

$$\begin{aligned}
c(b - a)^{-1}(x + b) + (x + a) &= c(b - a)^{-1}(x + a + b - a) + (x + a) \\
&= c + \underbrace{c(b - a)^{-1}(x + a)}_{\in (x+a)} + (x + a) \\
&= c + (x + a)
\end{aligned}$$

Similarly,

$$d(a - b)^{-1}(x + a) + (x + b) = d + (x + b)$$

Therefore,

$$\varphi(f(x)) = (c + (x + a), d + (x + b)),$$

as required.



Now onto its kernel. Let  $K = \ker \varphi$ . Let  $f(x) \in K$ . Then  $f(x) + (x + a) = 0$ , so  $f(x) \in (x + a)$  and  $f(x) = g_1(x)(x + a)$ , for some  $g_1(x) \in \mathbb{F}_p[x]$ . Thus,  $f(a) = 0$ .

Similarly,  $f(x) + (x + b) = 0$ , from which we get  $f(b) = 0$ .

Divide  $f(x)$  by  $(x + a)(x + b)$  in  $\mathbb{F}_p[x]$  (the leading coefficient of  $(x + a)(x + b)$  is 1, which is a unit in  $\mathbb{F}_p[x]$ , so the division is possible):

$$f(x) = q(x)(x + a)(x + b) + r(x),$$

where  $q(x) \in \mathbb{F}_p[x]$ , and  $r(x) \in \mathbb{F}_p[x]$  is of degree 1 or less (since  $(x + a)(x + b)$  is of degree 2). Furthermore,

$$0 = f(a) = r(a)$$

$$0 = f(b) = r(b)$$

So,  $r(x)$  is of degree 1 or less, and has two different roots, so we must have that  $r(x)$  is a zero polynomial:  $r(x) = 0$ . Thus,

$$f(x) = q(x)(x + a)(x + b) \in ((x + a)(x + b)) = (x^2 + 1)$$

Therefore,

$$K \subseteq (x^2 + 1)$$

Now let  $f(x) \in (x^2 + 1)$ . Then

$$f(x) = g(x)(x^2 + 1) = g(x)(x + a)(x + b),$$

for some  $g(x) \in \mathbb{F}_p[x]$ . But now it is clear that  $f(x) \in (x + a) \cap (x + b)$ , so

$$\varphi(f(x)) = ((x + a), (x + b)),$$

which is a zero in  $\mathbb{F}_p[x]/(x + a) \times \mathbb{F}_p[x]/(x + b)$ . Thus,  $f(x) \in K$ , so

$$K = (x^2 + 1)$$

Finally, using the First Isomorphism Theorem,

$$\mathbb{F}_p[x]/(x^2 + 1) \approx \mathbb{F}_p[x]/(x + a) \times \mathbb{F}_p[x]/(x + b)$$

Furthermore, by the First Isomorphism Theorem it is easy to see that the mapping  $\mathbb{F}_p \rightarrow \mathbb{F}_p$ ,  $x \rightarrow -a$  induces the isomorphism  $\mathbb{F}_p[x]/(x + a) \rightarrow \mathbb{F}_p$ . Similarly,  $\mathbb{F}_p[x]/(x + b) \approx \mathbb{F}_p$ .

Finally,

$$\mathbb{Z}[i]/(p) \approx \mathbb{F}_p[x]/(x^2 + 1) \approx \mathbb{F}_p[x]/(x + a) \times \mathbb{F}_p[x]/(x + b) \approx \mathbb{F}_p \times \mathbb{F}_p$$

## Result

If  $p$  is congruent to 3 modulo 4, then  $\mathbb{Z}[i]/(p)$  is a field isomorphic to  $\mathbb{F}_{p^2}$ .

If  $p = 2$ , then  $\mathbb{Z}[i]/(p) \approx \mathbb{F}_p[x]/(x^2 + 1)$ .

If  $p$  is congruent to 1 modulo 4, then  $\mathbb{Z}[i]/(p) \approx \mathbb{F}_p \times \mathbb{F}_p$ .

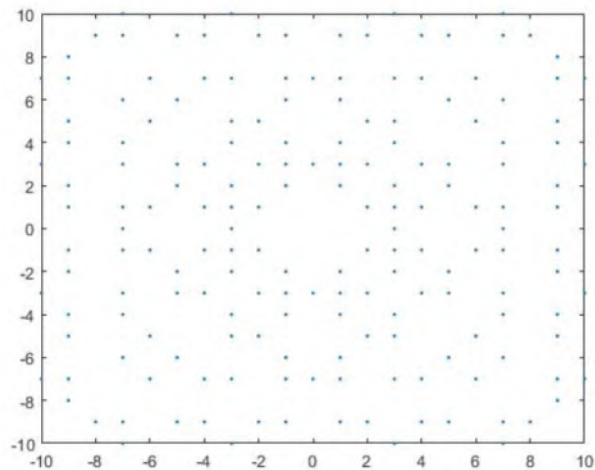
Let  $x \in \mathbb{Z}[\omega]$ ,  $x = a + b\omega$ . Then

$$x\bar{x} = a^2 - ab + b^2$$

Now it is easy to see that the similar result as **Theorem 12.5.2** holds; we only need to change  $a^2 + b^2$  to  $a^2 - ab + b^2$ .

Therefore, now we conclude that  $x = a + b\omega$  is a prime if and only if  $a \neq 0$ ,  $b \neq 0$ , and  $a^2 - ab + b^2$  is a prime integer, or  $a = 0$ ,  $b$  is a prime integer congruent to 3 modulo 4, or  $b = 0$ ,  $a$  is a prime integer congruent to 3 modulo 4.

Therefore, all primes in  $\mathbb{Z}[\omega]$  such that  $\sqrt{a^2 - ab + b^2} \leq 10$  are given by this picture:



## Result

2 of 2

Hint: modify **Theorem 12.5.2** to conclude that  $x = a + b\omega$  is a prime if and only if  $a \neq 0$ ,  $b \neq 0$ , and  $a^2 - ab + b^2$  is a prime integer, or  $a = 0$ ,  $b$  is a prime integer congruent to 3 modulo 4, or  $b = 0$ ,  $a$  is a prime integer congruent to 3 modulo 4.

9. a

(a)

First notice that

$$x^3 - 1 = (x^2 + x + 1)(x - 1)$$

Notice the following:  $x^2 + x + 1$  has a root in  $\mathbb{F}_p$  if and only if  $x^3 - 1$  has a root  $x_0 \neq 1$ . Truly, let  $x_0$  be a root of  $x^2 + x + 1$ , so  $x_0^2 + x_0 + 1 = 0$ . Then  $x_0^3 - 1 = 0$ , so it is a root of  $x^3 - 1$ . Furthermore,  $x_0 \neq 1$ , since  $1^2 + 1 + 1 = 3 \neq 0$  when  $p \neq 3$ .

On the other hand, suppose that  $x^3 - 1$  has a root  $x_0 \neq 1$ . Then  $(x_0^2 + x_0 + 1)(x_0 - 1) = 0$ . Since  $x_0 \neq 1$ , we must have that  $x_0^2 + x_0 + 1 = 0$ , meaning that  $x^2 + x + 1$  has a root.

Now notice the following: if  $x_0^3 - 1 = 0$ , that is,  $x_0^3 = 1$ , and  $x_0 \neq 1$ , then  $x_0$  is of order 3 in the multiplicative group  $\mathbb{F}_p^\times$ . Truly, let  $a$  be the order of  $x_0$ . Then  $a$  must divide 3 since  $x_0^3 = 1$ . This means that  $a = 1$  or  $a = 3$ . But  $a = 1$  yields  $x_0 = 1$ , which is a contradiction!

On the other hand, if  $\mathbb{F}_p^\times$  has an element  $y$  of order 3, then  $y^3 = 1$ ,  $y \neq 1$ , and  $y^3 - 1 = 0$ , so  $x^3 - 1$  has a root.

Thus:  $x^3 - 1$  has a root  $x_0 \neq 1$  if and only if the multiplicative group  $\mathbb{F}_p^\times$  has an element of order 3.

If  $\mathbb{F}_p^\times$  has an element of order 3, then 3 must divide the order of  $\mathbb{F}_p^\times$ . Since the order of  $\mathbb{F}_p^\times$  is  $|\mathbb{F}_p^\times| = p - 1$ , we conclude that 3 divides  $p - 1$ . But this means precisely that  $p \equiv 1$  modulo 3.

On the other hand, if  $p \equiv 1$  modulo 3, then  $p - 1$  is a multiple of 3, so the order of the group  $\mathbb{F}_p^\times$  is a multiple of a prime number 3. By **the Cauchy Theorem**, we conclude that there must exist some element  $x_0 \in \mathbb{F}_p^\times$  of order 3.

Thus:  $\mathbb{F}_p^\times$  has an element of order 3 if and only if  $p \equiv 1$  modulo 3.

Finally, linking everything together, we conclude that  $x^2 + x + 1$  has a root in  $\mathbb{F}_p$  if and only if  $p \equiv 1$  modulo 3, as required.

(b)

We will first prove that

$$\mathbb{Z}[x]/(x^2 + x + 1) \approx \mathbb{Z}[\omega]$$

To prove this, consider the mapping  $\mathbb{Z}[x] \rightarrow \mathbb{Z}[\omega]$  defined by  $x \rightarrow \omega$ . It is clearly a surjective homomorphism. Thus, we only need to prove that its kernel is  $(x^2 + x + 1)$ .

Denote by  $K$  the kernel of said mapping, and let  $f(x) \in K$ . Then  $f(\omega) = 0$ , meaning that  $\omega$  is a root of  $f(x)$ . Since  $f(x)$  has real coefficients, by the complex conjugate theorem we conclude that  $f(\bar{\omega}) = 0$ .

Now divide  $f(x)$  by  $x^2 + x + 1$  (we can do that in  $\mathbb{Z}[x]$  since  $x^2 + x + 1$  has 1 as the leading coefficient, which is a unit in  $\mathbb{Z}$ ):

$$f(x) = q(x)(x^2 + x + 1) + r(x),$$

where  $q(x), r(x) \in \mathbb{Z}[x]$ , and  $r(x)$  is of degree 1 or less.

Now it is easy to check that  $\omega$  and  $\bar{\omega}$  are roots of  $x^2 + x + 1$ . Therefore,

$$0 = f(\omega) = r(\omega)$$

$$0 = f(\bar{\omega}) = r(\bar{\omega})$$

Therefore,  $r(x)$  is a polynomial of degree 1 or less with 2 different roots. This means that  $r(x)$  is a zero polynomial; that is,  $r(x) = 0$ . Finally,

$$f(x) = q(x)(x^2 + x + 1) \in (x^2 + x + 1),$$

so

$$K \subseteq (x^2 + x + 1)$$

On the other hand, if  $g(x) \in (x^2 + x + 1)$ , then  $g(x) = h(x)(x^2 + x + 1)$  for some  $h(x) \in \mathbb{Z}[x]$ , so it is easy to see that  $g(\omega) = 0$ , meaning that  $g(x) \in K$ . Therefore,

$$K = (x^2 + x + 1)$$

By the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(x^2 + x + 1) \approx \mathbb{Z}[\omega]$$

We will consider the ring  $\overline{R} = R/(p)$ , and the following diagram:

$$\begin{array}{ccc} \mathbb{Z}[x] & \xrightarrow[\text{kill } p]{\text{kill}} & \mathbb{F}_p[x] \\ \downarrow \text{kill } x^2+x+1 & & \downarrow \text{kill } x^2+x+1 \\ R & \xrightarrow[\text{kill } p]{} & \overline{R} \end{array}$$

Therefore,  $(p)$  is a maximal ideal of  $R$  if and only if  $\overline{R}$  is a field by the **Proposition 11.8.2 (b)**. Using the above diagram, more precisely the fact that  $\mathbb{F}_p[x]/(x^2 + x + 1) \approx \overline{R}$ , this holds (by **Proposition 11.8.2 (b)**) if and only if  $(x^2 + x + 1)$  is a maximal ideal of  $\mathbb{F}_p[x]$ . Since  $\mathbb{F}_p[x]$  is a principal ideal domain, by **Corollary 12.2.9 (c)** we conclude that  $(x^2 + x + 1)$  is a maximal ideal of  $\mathbb{F}_p[x]$  if and only if  $x^2 + x + 1$  is irreducible in  $\mathbb{F}_p[x]$ .

Suppose that  $x^2 + x + 1$  has a root  $a$  in  $\mathbb{F}_p$ . Then we know (**Corollary 11.2.11**) that  $x - a$  divides  $x^2 + x + 1$ , so it is not irreducible.

On the other hand, if  $x^2 + x + 1$  is not irreducible, then there exists a polynomial  $f(x) \in \mathbb{F}_p[x]$  which divides  $x^2 + x + 1$ , and it is not a unit, nor is it associated with  $x^2 + x + 1$ . Thus, there exists a polynomial  $g(x) \in \mathbb{F}_p[x]$  such that

$$x^2 + x + 1 = f(x)g(x)$$

If  $f(x)$  is constant, then  $f(x) \neq 0$ , and it is a unit.

If  $g(x)$  is constant, then  $g(x) \neq 0$ , and it is a unit, so  $f(x)$  and  $x^2 + x + 1$  are associated.

Thus, neither  $f(x)$  nor  $g(x)$  are constant. Since their product is of degree 2, they must both be of degree 1. Thus,  $f(x) = ax + b$ , for some  $a, b \in \mathbb{F}_p$ ,  $a \neq 0$  (since it is of degree 1). Furthermore,  $f(-a^{-1}b) = 0$ . Thus,  $f(x)$  has a root in  $\mathbb{F}_p$ . But now  $x^2 + x + 1$  also has a root  $-a^{-1}b \in \mathbb{F}_p$ !

Therefore,  $x^2 + x + 1$  is irreducible in  $\mathbb{F}_p[x]$  if and only if it has no roots in  $\mathbb{F}_p$ . By part (a) of this exercise, this is equivalent to  $p \not\equiv 1$  modulo 3. Therefore,  $p \equiv 0$  modulo 3 or  $p \equiv 2 \equiv -1$  modulo 3. Furthermore,  $p \equiv 0$  modulo 3 means that  $p = 3k$ , so the only prime  $p$  such that  $p \equiv 0$  modulo 3 is  $p = 3$ . By the statement of the exercise,  $p \neq 3$ , so finally we conclude that  $x^2 + x + 1$  has no roots in  $\mathbb{F}_p$  if and only if  $p \equiv -1$  modulo 3.

Linking everything together, we have proven that  $(p)$  is a maximal ideal of  $R$  if and only if  $p \equiv -1$  modulo 3, as required.



(c)

Suppose that  $p$  factors in  $R$ . Since  $p$  is an integer prime, it is not a unit in  $R$ . Thus, since  $R$  is a unique factorization domain, there exists an irreducible  $\pi \in R$  which divides  $p$ . Furthermore, now  $\pi$  divides  $\bar{p} = p$ , so  $\pi\bar{\pi}$  divides  $p^2$ . However,  $\pi\bar{\pi} \in \mathbb{Z}$ , where  $p$  is prime, so  $\pi\bar{\pi} \in \{1, p, p^2\}$ . However,  $\pi\bar{\pi} = 1$  would mean that  $\pi$  is a unit, which is a contradiction (with it being irreducible).

Now suppose that  $\pi\bar{\pi} = p^2$ . Since  $\mathbb{Z}[\omega]$  is a unique factorization domain, more precisely from the uniqueness of the factorization, we conclude that  $p$  is then associated with  $\pi$ . But this also means that  $p$  is irreducible, which is impossible since it factors in  $R$ !

Therefore,  $p = \pi\bar{\pi}$ . Setting  $\pi = a - b\omega$ , we get

$$\bar{\pi} = a - b\bar{\omega} = a - be^{-2\pi i/3}$$

Thus,

$$\omega\bar{\omega} = e^{2\pi i/3}e^{-2\pi i/3} = 1$$

and

$$\omega + \bar{\omega} = e^{2\pi i/3} + e^{-2\pi i/3} = \cos(2\pi/3) + i\sin(2\pi/3) + \cos(-2\pi/3) + i\sin(-2\pi/3) = -1$$

Finally,

$$p = \pi\bar{\pi} = (a - b\omega)(a - b\bar{\omega}) = a^2 - ab(\omega + \bar{\omega}) + b^2 = a^2 + ab + b^2$$

To prove that the converse holds, suppose that  $p = a^2 + ab + b^2$ . Then just notice that

$$(a - b\omega)(a - b\omega^2) = a^2 - ab(\omega + \omega^2) + b^2\omega^3$$

Furthermore,

$$\omega^3 = 1$$

$$\omega + \omega^2 = e^{2\pi i/3} + e^{4\pi i/3} = \cos(2\pi/3) + i\sin(2\pi/3) + \cos(4\pi/3) + i\sin(4\pi/3) = -1$$

Therefore,

$$p = a^2 + ab + b^2 = (a - b\omega)(a - b\omega^2),$$

so it factors in  $R$ .

## Result

5 of 5

For (a), adopt the proof of **Lemma 12.5.5**.

For (b), adopt the proof of **Lemma 12.5.3**.

For (c), adopt the proof of **Theorem 12.5.2 (d)**.

10. a

(a)

We first write

$$\alpha = \pi_1^{a_1} \cdots \pi_n^{a_n}, \quad (1)$$

where  $\pi_i$  are Gauss primes such that  $\pi_i$  and  $\pi_j$  are not associated. Notice that

$$\bar{\alpha} = \bar{\pi}_1^{-a_1} \cdots \bar{\pi}_n^{-a_n}$$

and

$$\alpha \bar{\alpha} = (\pi_1 \bar{\pi}_1)^{a_1} \cdots (\pi_n \bar{\pi}_n)^{a_n}$$

Furthermore,  $\alpha \bar{\alpha}$  is a square integer, so

$$(\pi_1 \bar{\pi}_1)^{a_1} \cdots (\pi_n \bar{\pi}_n)^{a_n} = m^2$$

for some  $m \in \mathbb{Z}$ .

By **Theorem 12.5.2 (a)**, for each  $i = 1, 2, \dots, n$ ,  $\pi_i \bar{\pi}_i$  is either a prime integer or a square of a prime integer.

Suppose that  $\pi_j \bar{\pi}_j$  is a square of a prime integer for some  $j$ . Then

$$\pi_j \bar{\pi}_j = p^2$$

By the uniqueness of this factorization, we conclude that  $p$  is associated with  $\pi_j$ . Thus,  $\pi_j = vp$  for some unit  $v$ .

But now

$$\alpha = p(v\pi_1 \cdots \pi_{j-1}\pi_{j+1} \cdots \pi_n),$$

so  $p$  divides  $\alpha$ , which contradicts the statement of the exercise, since  $\alpha$  has no integer divisors!

Therefore,  $\pi_i \bar{\pi}_i = p_i$ , where  $p_i$  is a prime integer, for each  $i = 1, 2, \dots, n$ . Furthermore, if  $p_i = p_j$ , for  $i \neq j$ , then  $\pi_i \bar{\pi}_i = \pi_j \bar{\pi}_j$ , so, once again, by the uniqueness of this factorization,  $\pi_i$  and  $\pi_j$  are associated. But this contradicts the properties of the factorization (1)! Therefore,  $p_i \neq p_j$ , for all  $i, j \in \{1, \dots, n\}$ ,  $i \neq j$ .

Now  $\alpha \bar{\alpha} = m^2$  becomes

$$p_1^{a_1} \cdots p_n^{a_n} = m^2$$

Since  $p_i$  are prime integers, for their product to be a square of an integer, we must have that 2 divides  $a_i$ ; that is,  $a_i = 2b_i$ , for some  $b_i$ .

Finally, (1) now becomes

$$\alpha = (\pi_1^{b_1} \cdots \pi_n^{b_n})^2,$$

so  $\alpha$  is a square in  $\mathbb{Z}[i]$ .



(b)

(I assume that  $c$  is a positive integer, since otherwise  $c = m^2 + n^2$  is absurd.)

Let  $\alpha = a + bi$ . Suppose that  $d \in \mathbb{Z}$  divides  $\alpha$ . Then, by **Lemma 12.5.1**  $d$  divides both  $a$  and  $b$ . Since  $a$  and  $b$  are relatively prime, this implies that  $d = \pm 1$ . But this means that  $\alpha$  has no nontrivial integer factors!

Furthermore,

$$\alpha\bar{\alpha} = (a + bi)(a - bi) = a^2 + b^2 = c^2$$

Now using (a),  $\alpha$  is a square in  $\mathbb{Z}[i]$ ; meaning, there exist  $m, n \in \mathbb{Z}$  such that

$$\alpha = (m + ni)^2 = m^2 - n^2 + 2mni \implies a + bi = m^2 - n^2 + 2mni$$

Therefore,

$$a = m^2 - n^2$$

$$b = 2mn$$

Furthermore,

$$c^2 = a^2 + b^2 = m^4 + n^4 - 2m^2n^2 + 4m^2n^2 = (m^2 + n^2)^2$$

Thus,  $c = m^2 + n^2$ .

### Result

(a) Use the fact that  $\mathbb{Z}[i]$  has the property of prime factorization, and **Theorem 12.5.2 (a)**.

(b) Use the (a) part, with  $\alpha = a + bi$ .

## Miscellaneous Problem

1. a

Suppose that in  $S$  the following holds:

$$ab = 0 \implies a = 0 \text{ or } b = 0 \quad (1)$$

Also, we say that *factoring terminates* if each  $a$  can be factored a finite number of times:

$$a = b_1 a_1 = b_1 b_2 a_2 = \dots = b_1 b_2 \dots b_n,$$

where  $b_i$  are such that  $b_i = c_1 c_2$  implies  $c_1 = 1$  or  $c_2 = 1$ .

We say that  $c$  is *irreducible* if  $c = d_1 d_2$  implies  $d_1 = 1$  or  $d_2 = 1$ .

We say that  $p$  is *prime* if  $ab = pc$  implies  $a = pd$  or  $b = pe$ , where  $d, e \in S$ .

We say that  $u$  is a *unit* if  $u$  is invertible.

We say that  $S$  is a *unique factorization domain* if factoring terminates and if

$$a = b_1 b_2 \dots b_n$$

and

$$a = c_1 c_2 \dots c_m$$

implies that  $b_i = uc_j$  for some unit  $u$  and some  $j \in \{1, \dots, m\}$ , for every  $i = 1, \dots, n$ , and vice versa ( $c_i = vb_j$ , for some unit  $v$ ).

Finally, onto the statement:

**Proposition.** Suppose that (1) holds in  $S$ , and that factoring terminates. Then  $S$  is a unique factorization domain if and only if every irreducible element of  $S$  is also prime.

## Result

2 of 2

$$ab = 0 \implies a = 0 \text{ or } b = 0 \quad (1)$$

**Proposition.** Suppose that (1) holds in  $S$ , and that factoring terminates. Then  $S$  is a unique factorization domain if and only if every irreducible element of  $S$  is also prime.

2. a

(a)

To determine that whether the semi group  $S$  has unique factorization when coordinates of the vectors  $v_i$  are nonnegative,

Suppose  $S$  be the semi-group with identity element with law of composition addition.

Then, for  $v_1, v_2, \dots, v_n \in \mathbb{Z}^2$ ,  $a_i$  be the nonnegative integer coefficients and  $S$  is the semi group of all combinations such that,

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n$$

Suppose for any  $v_i \in \mathbb{Z}^2$  such that

$$v_i = (x_i, y_i)$$

Then,

$$\begin{aligned} a_1 v_1 + a_2 v_2 + \dots + a_n v_n &= a_1 (x_1, y_1) + a_2 (x_2, y_2) + \dots + a_n (x_n, y_n) \\ &= (a_1 x_1 + a_2 x_2 + \dots + a_n x_n, a_1 y_1 + a_2 y_2 + \dots + a_n y_n) \end{aligned}$$

Since, the coordinates of the vectors  $v_i$  are nonnegative and  $a_i$  are nonnegative integer coefficients.

Then,

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n \geq 0, a_1 y_1 + a_2 y_2 + \dots + a_n y_n \geq 0$$

Then, the coordinates  $a_1 x_1 + a_2 x_2 + \dots + a_n x_n$  and  $a_1 y_1 + a_2 y_2 + \dots + a_n y_n$  of the elements of the semi group  $S$  are nonnegative.

Then, for  $x_i \neq y_i$ , the vectors of the semi group  $S$  will have the distinct and unique prime factors for each of its vectors.

Then, each of the elements of the semi-group  $S$  will have unique factorization.

Hence, **the semi-group  $S$  will have unique factorization whenever coordinates of the vectors  $v_i$  are nonnegative.**

---

(b)

To determine that whether the semi group  $S$  has unique factorization when coordinates of the vectors  $v_i$  are any integer,

Suppose  $S$  be the semi-group with identity element with law of composition addition.

Then, for  $v_1, v_2, \dots, v_n \in \mathbb{Z}^2$ ,  $a_i$  be any integer coefficients and  $S$  is the semi group of all combinations such that,

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n$$

Suppose for any  $v_i \in \mathbb{Z}^2$  such that

$$v_i = (x_i, y_i)$$

Then,

$$\begin{aligned} a_1 v_1 + a_2 v_2 + \dots + a_n v_n &= a_1 (x_1, y_1) + a_2 (x_2, y_2) + \dots + a_n (x_n, y_n) \\ &= (a_1 x_1 + a_2 x_2 + \dots + a_n x_n, a_1 y_1 + a_2 y_2 + \dots + a_n y_n) \end{aligned}$$

Since, the coordinates of the vectors  $v_i$  are any integer and  $a_i$  are nonnegative integer coefficients.

Then, the coordinates of the semi group  $S$  may be any integer.

Then, there arises two cases,

Case-1, when  $x_i = y_i$

Then, both the coordinates of the semi group  $S$  will have the same factor.

Then, in this case, elements of the semi group  $S$  will not have unique factorization.

Case-2, when  $x_i \neq y_i$

Then, the one of the coordinate of element of  $S$  may be multiple of the other coordinate as  $x_i$  and  $y_i$  are the any integer.

Then, both the coordinates of the semi group  $S$  may or may not have the same factor.

Then, each of the elements of the semi-group  $S$  may or may not have unique factorization.

Hence, **the semi-group  $S$  may or may not have unique factorization whenever coordinates of the vectors  $v_i$  are any integers.**

### 3. a

Consider the provided statement to prove that  $n \geq p-1$  where  $A$  is matrix of order  $n \times n$  such that  $A^p = I$  and  $p$  be an integer prime.

[Comment](#)

Step 2 of 4 ^

It is consider that  $A$  to be an operator on a  $n$  dimensional complex vector space  $V = \mathbb{C}^n$ .

For vectors  $v = (v_1, \dots, v_n)$  and  $w = (w_1, \dots, w_n)$ . It is defined that,  $\langle v, w \rangle_0 = \bar{v}_1 w_1 + \dots + \bar{v}_n w_n$  to be the standard complex dot product of  $v$  and  $w$  such that,

$$\begin{aligned}\langle v, v \rangle &= \sum_{i=1}^n |v_i|^2 \\ &= |v|^2 > 0\end{aligned}$$

It is assumed that,

$$\langle v, w \rangle = \sum_{i=0}^{k-1} \langle A^i v, A^i w \rangle_0$$

It can be claim that  $\langle \cdot, \cdot \rangle$  is a positive definite Hermetian form and clearly  $\langle \cdot, \cdot \rangle_0$  is a Hermetian form and for non-zero  $v \in V$ ,

$$\begin{aligned}\langle v, v \rangle &= \sum_{i=0}^{k-1} |A^i v|^2 \\ &> 0\end{aligned}$$

For any  $v, w \in V$ ,  $\langle Av, Aw \rangle = \langle v, w \rangle$ . Therefore,  $A$  is unitary and so  $AA^* = I_n = A^*A$  and  $A$  is normal. So, from the Artin theorem  $A$  is diagonalizable.

Let  $c(x) \in \mathbb{Z}[x]$  be the characteristic polynomial of  $A$  and also considering  $A$  is as a complex matrix. As it is given that,  $A \neq I_n$  is at least one diagonal entry that is eigenvalue  $\lambda$  which is not equal to 1.

It is assumed that  $v$  is a corresponding eigenvector therefore  $Av = \lambda v$ . As,

$$\begin{aligned}v &= A^p v \\ &= \lambda^p v\end{aligned}$$

Therefore,  $\lambda$  is a nontrivial root of  $x^p - 1$ , it means that  $\lambda$  is a root of the cyclotomic polynomial  $f_p(x) = x^{p-1} + \dots + x + 1$ .

But as  $\lambda$  is an eigenvalue of  $A$  and it is also given that  $c(\lambda) = 0$ . Since  $f_p(x)$  is a primitive irreducible polynomial in  $\mathbb{Z}[x]$ ,  $f_p$  divides  $c$  and  $n = \deg c \geq p-1$ .

Hence, provided statement is **proved**.

Here is an example such with  $n = p-1$ . It is assumed that  $p = 2$  which is a prime integer.

As it is given that  $A \neq I$  then,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Now calculate,

$$\begin{aligned}A^2 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1+0 & 0+0 \\ 0+0 & (-1)(-1) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= I\end{aligned}$$

### 4. a

First Isomorphism Theorem- Consider a group  $G$  and let  $\phi : G \rightarrow Y$  be a homomorphism. Then image of  $\phi$  is isomorphic to the quotient group  $G/\ker \phi$ .

Mathematically,

$$\frac{G}{\ker \phi} \cong \phi(G)$$

Laurent Ring-The ring of all the polynomials over a field say  $F$  with variables  $x$  and  $x^{-1}$ . It is denoted by  $F[x, x^{-1}]$ .

(a)

Let  $R$  be the ring of functions that are polynomials in  $\cos t$  and  $\sin t$ .

Define  $x(t) = \cos t$  and  $y(t) = \sin t$ .

Now consider the ring of polynomials  $R[x, y]$ .

Consider a map  $\phi : R[x, y] \rightarrow R[t]$  defined by

$$\phi(g(x, y)) = g(\cos t, \sin t)$$

This map is a homomorphism since it is defined on polynomials.

Evaluate kernel of this map

Since,

$$\cos^2 t + \sin^2 t = 1$$

$$\cos^2 t + \sin^2 t - 1 = 0$$

Define a polynomial  $g(x, y) \in R[x, y]$  by

$$g(x, y) = x^2 + y^2 - 1$$

Then, clearly  $g(x, y) \in \ker \phi$

So  $\langle g(x, y) \rangle \subseteq \ker \phi$

Let  $f(x, y) \in \ker \phi$

Consider  $f$  as a polynomial with variable term as  $x$  only and coefficients from  $R[y]$ .

$$\text{Then } f(x) = (x^2 + y^2 - 1)h(x) + r(x), \deg r(x) \leq 1$$

This implies that  $r(x)$  is a linear polynomial, so there exist coefficients  $a, b \in R[y]$  which satisfy the below mentioned relation

$$r(x) = a + bx$$

Also since  $f(x, y) \in \ker \phi$ , so  $f(\cos t, \sin t) = 0$

Thus  $r(\cos t, \sin t) = 0$

This implies  $a(\sin t) + b(\sin t)\cos t = 0$ .

Squaring both sides and rearranging the terms

$$(a(\sin t))^2 = (-b(\sin t)\cos t)^2$$

$$a^2(\sin t) = b^2(\sin t)\cos^2 t$$

Use the trigonometric identity  $\cos^2 t = 1 - \sin^2 t$

$$a^2(\sin t) = b^2(\sin t)(1 - \sin^2 t)$$

Since,  $a, b \in R[y]$ , means they both are polynomials in  $R[y]$

So,

$$a^2(\sin t) = b^2(\sin t)(1 - \sin^2 t)$$

$$a^2(y) = b^2(y)(1 - y^2)$$

From above equation there is no element of power 2 on the left hand side. So  $b = 0$  and  $a = 0$

Thus any element  $f(x, y) \in \ker \phi$  is entirely divisible by  $g(x, y) = x^2 + y^2 - 1$ .

So  $\ker \phi = \langle x^2 + y^2 - 1 \rangle$

Now by First Isomorphism Theorem,

$$\frac{R[x, y]}{\ker \phi} \cong \phi(R[x, y])$$

$$\frac{R[x, y]}{\ker \phi} \cong R[t]$$

Therefore, for any ring  $R$  of functions consisting of polynomial in  $\cos t$  and  $\sin t$  with real coefficients the following holds  $(R[x, y]/\ker \phi) \cong R[t]$ .

(b)

Consider a polynomial  $f(t) = \sin^2 t$  in  $R[t]$ .

There are two ways of factorizing this polynomial in  $R[t]$  which are as follows

$$\begin{aligned} \sin^2 t &= (\sin t)(\sin t) \\ \sin^2 t &= 1 - \cos^2 t \\ &= (1 - \cos t)(1 + \cos t) \end{aligned}$$

$$\text{Hence, } (\sin t)(\sin t) = \sin^2 t = (1 - \cos t)(1 + \cos t)$$

Since the factorization of a polynomial in  $R[t]$  is not unique.

Thus,  $R$  is not a UFD.

Therefore, for any ring  $R$  of functions consisting of polynomial in  $\cos t$  and  $\sin t$  with real coefficients is not a UFD.

(c)

$$\text{Let } S = \mathbb{C}[x, y] / (x^2 + y^2 - 1)$$

Let  $u, v$  be variables in the polynomial ring over complex

$$\text{If } uv - 1 = 0, \text{ then } uv = 1$$

This implies that  $u, v$  are inverse of each other.

$$\text{So } v = u^{-1}$$

Define Laurent polynomial ring as

$$\frac{\mathbb{C}[u, v]}{(uv - 1)}$$

Define a map  $\phi: \frac{\mathbb{C}[u, v]}{(uv - 1)} \rightarrow S$  by

$$\phi(f(u, v)) = f(x + iy, x - iy) = f(x, y)$$

Now let  $f_1, f_2 \in \frac{\mathbb{C}[u, v]}{(uv - 1)}$  be arbitrary elements such that they attain the same value when operated by  $\phi$ .

Then,

$$\begin{aligned} \phi(f_1(u, v)) &= \phi(f_2(u, v)) \\ f_1(x + iy, x - iy) &= f_2(x + iy, x - iy) \\ f_1(u, v) &= f_2(u, v) \end{aligned}$$



Hence the above defined map is injective.

Also for every element  $g(x, y) \in S$  define  $x = (u+v)/2$  and  $y = (u-v)/2i$

Then, choose  $h((u+v)/2, (u-v)/2i) \in$  Laurent polynomial ring such that

$$\phi(h((u+v)/2, (u-v)/2i)) = h(x, y)$$

Hence the map is surjective also.

Let  $f_1, f_2 \in \mathbb{C}[u, v]/(uv-1)$ .

Consider

$$\begin{aligned}\phi((f_1 + f_2)(u, v)) &= f_1 + f_2(x + iy, x - iy) \\ &= f_1 + f_2(x, y) \\ &= f_1(x, y) + f_2(x, y) \\ &= \phi(f_1) + \phi(f_2)\end{aligned}$$

Hence the map is a bijective homomorphism thus an isomorphism.

So

$$\frac{\mathbb{C}[u, v]}{(uv-1)} \cong S$$

Since Laurent polynomial ring is a ring of fractions with operation as multiplication.

Hence  $S$  is a UFD, moreover a principal ideal domain.

**Therefore, the ring  $S = \mathbb{C}[x, y]/(x^2 + y^2 - 1)$  is a PID as well as a unique factorization domain.**

(d)

Use the result of part (a) for every element  $f(x, y) \in R[x, y]$  can be written as

$$f = a(y) + b(y)x$$

Define a function  $\sigma: R[x, y] \rightarrow R[y]$  by

$$\begin{aligned}\sigma(f) &= \sigma(a(y) + b(y)x) \\ &= a^2(y) + b^2(y)(y^2 - 1)\end{aligned}$$

Let  $f, g \in R[x, y]$  be arbitrary element

Then  $f = a(y) + b(y)x$  and  $g = c(y) + d(y)x$

Now,

$$\begin{aligned}fg &= (a(y) + b(y)x)(c(y) + d(y)x) \\ &= a(y)c(y) + (a(y)d(y) + b(y)c(y))x + b(y)d(y)x^2\end{aligned}$$

Operate  $\sigma: R[x, y] \rightarrow R[y]$  on  $fg$

$$\begin{aligned}\sigma(fg) &= \sigma(a(y)c(y) + (a(y)d(y) + b(y)c(y))x + b(y)d(y)x^2) \\ &= a^2(y)c^2(y) + a^2(y)d^2(y) + b^2(y)c^2(y)(y^2 - 1) + b^2(y)d^2(y)(y^2 - 1)^2 \\ &= (a^2(y) + b^2(y)(y^2 - 1))(c^2(y) + d^2(y)(y^2 - 1)) \\ &= \sigma(f)\sigma(g)\end{aligned}$$

Thus  $\sigma: R[x, y] \rightarrow R[y]$  is a multiplicative function.

So this implies that the units of ring  $R[x, y]$  are all the non-zero real numbers.

For  $S = \mathbb{C}[x, y]/(x^2 + y^2 - 1)$  clearly from the part(c) the invertible elements are simply the non-zero complex numbers along with the powers of  $x + iy$ ,  $x - iy$  and multiplication of these functions with the non-zero complex numbers.

**Therefore, the units of ring  $R[x, y]$  are all the non-zero real numbers and the units of the ring  $S = \mathbb{C}[x, y]/(x^2 + y^2 - 1)$  are non-zero complex numbers along with the powers of  $x + iy$ ,  $x - iy$  and multiplication of these functions with the non-zero complex numbers.**

**Statement:**  $x^2 + y^2 = n$  has a point with integer coefficients if and only if every prime  $p \equiv 3$  modulo 4 in the prime factorization of  $n$  has an even exponent.

Suppose that  $x^2 + y^2 = n$  has a point  $(x_0, y_0)$  with integer coefficients. Then

$$x_0^2 + y_0^2 = n \implies (x_0 + y_0i)(x_0 - y_0i) = n$$

Let  $p \equiv 3$  modulo 4 be a prime which divides  $n$ . Such  $p$  is also a Gauss prime by **Theorem 12.5.2 (c)**, so  $p$  divides either  $x_0 + y_0i$  or  $x_0 - y_0i$ . Suppose that it divides  $x_0 + y_0i$ . Then by **Lemma 12.5.1**  $p$  divides  $x_0$  and  $y_0$ . Again by the same Lemma, we conclude that  $p$  divides  $x_0 - y_0i$ . Similarly, if we assume that  $p$  divides  $x_0 - y_0i$ , we can easily prove that  $p$  divides  $x_0 + y_0i$ . Since  $n = (x_0 + y_0i)(x_0 - y_0i)$ ,  $p$  must have an even exponent in the prime decomposition of  $n$ . Truly, let  $n = ap^k$ , where  $a$  is an integer which  $p$  does not divide. Then  $p$  divides  $n$ . It also divides  $x_0 + y_0i$  and  $x_0 - y_0i$ , so  $x_0 + y_0i = p(b + ci)$  and  $x_0 - y_0i = p(d + ei)$ . Furthermore,

$$\begin{aligned} x_0 - y_0i &= \overline{x_0 + y_0i} \implies p(d + ei) = \underbrace{\bar{p}}_{=p} (b - ci) \\ &\implies p((d - b) + (e + c)i) = 0 \\ &\implies d - b + (e + c)i = 0 \\ &\implies d = b, e = -c \end{aligned}$$

Thus,  $x_0 - y_0i = p(b - ci)$ . Finally,

$$n = (x_0 + y_0i)(x_0 - y_0i) \implies ap^k = p^2(b + ci)(b - ci) \implies ap^{k-2} = (b + ci)(b - ci)$$

We can repeat this process, and in finitely many repetitions we will get either  $k - 2l = 0$  or  $k - 2l = 1$ , for some  $l \in \mathbb{Z}$ . When  $k - 2l = 1$ , we get, for some integers  $d, d', e, e' \in \mathbb{Z}$ ,

$$ap = (d + ei)(d - ei) \implies ap = p^2(d' + e'i)(d' - e'i) \implies a = p(d' + e'i)(d' - e'i)$$

Therefore,  $p$  divides  $a$ , which is a contradiction. Thus,  $k - 2l = 1$  is impossible, so  $k - 2l = 0$ , for some integer  $l$ , that is,  $k = 2l$ , so it is even.

Now suppose that  $n$  has the property given in the **Statement**. Write

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$$

Grouping those  $p_i$  with an odd exponent together, and those with an even exponent together, we can write

$$n = ab^2,$$

where  $a$  is square free. Also note that if  $p$  divides  $a$ , then  $p \not\equiv 3$  modulo 4 by construction of  $a$  and  $b$ . This means that, if  $p$  divides  $a$ , then  $p \equiv 1$  modulo 4 or  $p = 2$ . By **Theorem 12.5.2 (d)** this means that each such  $p$  can be written as  $p = \pi_p \bar{\pi}_p$ , for some Gauss prime  $\pi_p$ .

Let  $q_1, q_2, \dots, q_m$  be primes such that  $a = q_1 \cdots q_m$ . Define a complex number

$$z = \pi_{q_1} \cdots \pi_{q_m} b$$

If we write  $z = x_0 + y_0i$ , we conclude (from the above expression) that  $x_0$  and  $y_0$  are integers. Furthermore,

$$x_0^2 + y_0^2 = (x_0 + y_0i)(x_0 - y_0i) = \pi_{q_1} \bar{\pi}_{q_1} \cdots \pi_{q_m} \bar{\pi}_{q_m} \underbrace{b \bar{b}}_{=b} = q_1 \cdots q_m b^2 = ab^2 = n$$

## Result

3 of 3

**Statement:**  $x^2 + y^2 = n$  has a point with integer coefficients if and only if every prime  $p \equiv 3$  modulo 4 in the prime factorization of  $n$  has an even exponent.

Suppose that for some  $i \in \{1, \dots, r\}$  we have that

$$Q_j \not\subseteq P_i$$

for every  $j = 1, \dots, s$ . Recall the canonical projections

$$\pi_i : R \rightarrow R/P_i, \quad \pi_i(r) = r + P_i$$

Since  $R$  is a domain,  $R/P_i$  is also a domain. Furthermore,  $Q_j \not\subseteq P_i$  means that there exists some  $q_j \in Q_j$  such that  $q_j \notin P_i$ . This also means that

$$\pi_i(q_j) = q_j + P_i \neq P_i$$

Therefore, since  $R$  is a domain and  $\pi_i$  is a homomorphism,

$$\pi_i(q_1 \cdots q_s) = \pi_i(q_1) \cdots \pi_i(q_s) \neq P_i$$

Therefore,

$$q_1 \cdots q_s + P_i \neq P_i \implies q_1 \cdots q_s \notin P_i$$

However,

$$Q_1 \cdots Q_s = P_1 \cdots P_r \subseteq P_i,$$

which, since  $q_1 \cdots q_s \in Q_1 \cdots Q_s$ , implies  $q_1 \cdots q_s \in P_i$ . Contradiction!

Thus, there must exist  $j \in \{1, \dots, s\}$  such that  $Q_j \subseteq P_i$ . Since  $Q_j$  and  $P_i$  are maximal ideals, we also conclude that  $Q_j = P_i$ .

Therefore, every  $P_i$  is equal to some  $Q_j$ . Switching the roles of  $P$  and  $Q$  in the above proof yields the converse: every  $Q_j$  is equal to some  $P_i$ . This proves the statement of the exercise.

## Result

2 of 2

Hint to start: suppose that there exists some  $P_i$  such that  $Q_j \not\subseteq P_i$  for every  $j$ , and observe the canonical projections

$$\pi_i : R \rightarrow R/P_i, \quad \pi_i(r) = r + P_i$$

## 7. a

### (a)

Let it assume that  $M \subset R$  be a maximal ideal. From the previous exercise it is not principal therefore there exist  $f_1, f_2 \in M$  such that they do not share a common factor that is  $f_1, f_2$  do not share a common factor in  $\mathbb{Q}[x]$ .

From the theorem 12.3.6, thus  $r_0 f_1 + s_0 f_2 = 1$  for some  $r, s \in \mathbb{Q}[x]$  therefore denominators are  $r f_1 + s f_2 = q \in \mathbb{Q}$ . Where  $r, s \in \mathbb{Z}[x]$  that is  $M \cap \mathbb{Z} \neq (0)$

Now  $M \cap \mathbb{Z}$  is a prime, since if  $ab \in M \cap \mathbb{Z}$  then  $a \in M$  and moreover  $a \in \mathbb{Z}$  for otherwise  $ab \notin \mathbb{Z}$ . Thus,  $M \cap \mathbb{Z} = (0)$  or  $(p)$  for some prime so the above case is impossible.

Now it is consider that the image of  $M$  is  $M'$  in  ${}_p[x]$  then  $\frac{\mathbb{Z}[x]}{M} \approx \frac{F_p[x]}{M'}$  is a field.

So, from the proposition 11.8.2  $M'$  is maximal and from Cor. 12.2.9, it is generated by some irreducible  $f_0 \in F_p[x]$ . Since,  $F_p[x]$  is a PID by the proposition 12.2.5.

If  $f \in \mathbb{Z}[x]$  is a lift of  $f_0$  then  $(p, f) \subset M$ , as  $f = 0$  in  $\frac{F_p[x]}{M'}$  such that  $M \subset (p, f)$

because if  $g \in \frac{M}{(p, f)}$  then

$$\begin{aligned} g \neq 0 &\in \approx \frac{F_p[x]}{M'} \\ &\approx \frac{\mathbb{Z}[x]}{M} \end{aligned}$$

(b)

From previous part, if  $I = (f, g)$  and  $f, g$  have no common factors then  $I \cap \mathbb{Z} \neq (0)$ .

Since,  $\mathbb{Z}$  is a PID then  $M \cap \mathbb{Z} = n$  where  $n \in \mathbb{Z}$  and the value of  $n$  is not equal to  $\pm 1$  otherwise  $\frac{R}{I} = 0$ .

Therefore, it is assume that  $n = \prod p_i^{k_i}$  is a prime factorization. Then, it is claim that

$$\begin{aligned} \frac{R}{I} &\approx \frac{\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)[x]}{(f, g)} \\ &\approx \frac{\prod_i \left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x]}{(f, g)} \\ &\approx \prod_i \frac{\left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x]}{(f, g)} \end{aligned}$$

The first isomorphism is proved and from the exercise 11.6.8 second isomorphism is followed repeatedly. For the third, the canonical surjection is considered as below:

$$\pi: \prod_i \left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x] \rightarrow \prod_i \frac{\left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x]}{(f, g)}$$

Since, an element on the left maps to zero on the right if and only if each direct factor is in  $(f, g)$  therefore  $\ker \pi = (f, g)$ . So, it holds by the first isomorphism theorem.

Let  $R_i$  be the direct factor on the right. It is to be shown that each  $R_i$  is finite, since it is already define that without loss of generality  $p_i \nmid f$  since  $f, g$  do not share a common factor. Then,  $f$  can be written as  $f = f_1 - f_2$  where  $p_i \nmid f_1$  while  $p_i \mid f_2$ .

As,  $\left(\frac{\mathbb{Z}}{p_i^{k_i}}\right)[x]$  the  $p_i^{k_i} \mid f_2^{k_i}$

$$\begin{aligned} f_1^{k_i} &= f_1^{k_i} - f_2^{k_i} \\ &= (f_1 - f_2)h \\ &= fh \in (f, g) \end{aligned}$$

Where  $h \in \left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x]$  and if  $a$  is a leading coefficient of  $f_1$  then  $f_1^{k_i}$  has the leading coefficient  $a^{k_i}$ . This is a unit since if not then from the theorem 11.9.2,  $a^{k_i}$  is contained in a maximal ideal of  $\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}$ . But from the correspondence theorem  $p_i$  is the unique maximal ideal of this ring, therefore contradiction occurs that  $p_i \mid a$ .



Therefore  $m = a^{-k_i} f_i^{k_i}$  is a monic polynomial in  $\left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x]$  so each  $R_i$  is finite. As any polynomial whose degree is greater than or equal to  $\deg m$  can be reduced to a polynomial which has lower degree by using  $m = 0$  and there are only finitely many polynomials in  $\left(\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}\right)[x]$  has degree less than  $\deg m$  because  $\frac{\mathbb{Z}}{p_i^{k_i}\mathbb{Z}}$  is finite.

8. a

Let  $R' = \mathbb{Z}[\alpha]$ . By definition,

$$\mathbb{Z}[x]/(vx - u) \approx \mathbb{Z}[\alpha]$$

Now define a mapping  $\mathbb{Z}[x] \rightarrow \mathbb{Z}\left[\frac{u}{v}\right]$  by  $x \mapsto \frac{u}{v}$ . It is clearly a homomorphism. Moreover, recall that

$$\mathbb{Z}[\beta] = \{a_0 + a_1\beta + \dots + a_n\beta^n \mid n \in \mathbb{N} \cup \{0\}, a_i \in \mathbb{Z}\}, \quad (1)$$

therefore this mapping is also surjective. Thus, we only need to find its kernel, which we denote by  $K$ .

Let  $f(x) \in K$ . Then  $f\left(\frac{u}{v}\right) = 0$ . This means that  $\frac{u}{v}$  is a root of  $f(x)$ . In  $\mathbb{Q}[x]$ , we can divide  $f(x)$  by  $vx - u$  (since  $v$  is a unit in  $\mathbb{Q}$ ):

$$f(x) = q(x)(vx - u) + r(x),$$

where  $q(x), r(x) \in \mathbb{Q}[x]$ , and  $r(x)$  is constant (since  $vx - u$  is linear). Furthermore,

$$0 = f\left(\frac{u}{v}\right) = r\left(\frac{u}{v}\right),$$

so  $r(x) = 0$  (a zero polynomial). Therefore, in  $\mathbb{Q}[x]$ ,

$$f(x) = q(x)(vx - u)$$

Since  $vx - u$  is a primitive polynomial in  $\mathbb{Z}[x]$  (remember that  $u, v$  are relatively prime), and it divides an integer polynomial  $f(x)$  in  $\mathbb{Q}[x]$ , by **Theorem 12.3.6 (a)** we conclude that  $vx - u$  also divides  $f(x)$  in  $\mathbb{Z}[x]$ . Thus,

$$f(x) = \tilde{q}(x)(vx - u),$$

with some  $\tilde{q}(x) \in \mathbb{Z}[x]$ . Therefore,

$$f(x) \in (vx - u) \implies K \subseteq (vx - u)$$

For the other inclusion, let  $g(x) \in (vx - u)$ . Then  $g(x) = h(x)(vx - u)$ , for some  $h(x) \in \mathbb{Z}[x]$ . Now it is clear that

$$g\left(\frac{u}{v}\right) = 0,$$

so

$$g(x) \in (vx - u),$$

and

$$K = (vx - u)$$

By the First Isomorphism Theorem,

$$\mathbb{Z}[x]/(vx - u) \approx \mathbb{Z}\left[\frac{u}{v}\right]$$

Thus,

$$R' \approx \mathbb{Z}\left[\frac{u}{v}\right]$$

For the other part, we will prove that

$$\mathbb{Z}\left[\frac{u}{v}\right] = \mathbb{Z}\left[\frac{1}{v}\right]$$

Since  $\frac{u}{v} = u \cdot \frac{1}{v}$ , by the definition of  $\mathbb{Z}\left[\frac{1}{v}\right]$  (see (1)) we conclude that  $\frac{u}{v} \in \mathbb{Z}\left[\frac{1}{v}\right]$ . Since  $\mathbb{Z}\left[\frac{u}{v}\right]$  is the smallest ring which contains  $\frac{u}{v}$ , we conclude that

$$\mathbb{Z}\left[\frac{u}{v}\right] \subseteq \mathbb{Z}\left[\frac{1}{v}\right]$$

On the other hand, since  $u, v$  are relatively prime, there exist integers  $a, b$  such that

$$au + bv = 1$$

Dividing the above equation by  $v$ ,

$$b + a \cdot \frac{u}{v} = \frac{1}{v}$$

By the definition of  $\mathbb{Z}\left[\frac{u}{v}\right]$ ,  $b + a \cdot \frac{u}{v} \in \mathbb{Z}\left[\frac{u}{v}\right]$ , so  $\frac{1}{v} \in \mathbb{Z}\left[\frac{u}{v}\right]$ . Since  $\mathbb{Z}\left[\frac{1}{v}\right]$  is the smallest ring which contains  $\frac{1}{v}$ , we conclude that

$$\mathbb{Z}\left[\frac{1}{v}\right] \subseteq \mathbb{Z}\left[\frac{u}{v}\right]$$

Finally,

$$\mathbb{Z}\left[\frac{u}{v}\right] = \mathbb{Z}\left[\frac{1}{v}\right]$$

This also proves that

$$R' \approx \mathbb{Z}\left[\frac{1}{v}\right]$$

## Result

3 of 3

By definition,  $R' \approx R[x]/(vx - u)$ . Now use the First Isomorphism Theorem to prove that  $R' \approx \mathbb{Z}\left[\frac{u}{v}\right]$ .

For the second part, we can prove that  $\mathbb{Z}\left[\frac{1}{v}\right] = \mathbb{Z}\left[\frac{u}{v}\right]$ .

9. a



To explain the way to determine the index  $[V:W]$ ,

Suppose  $R = \mathbb{Z}[i]$  be the ring of Gauss integer and  $W$  be the  $R$ -submodule of  $V = R^2$  generated by columns of  $2 \times 2$  matrix,

Then,  $W$  is a vector space of  $V$  over  $R$ .

Then, for any  $x, y \in W, \alpha \in V$  such that

$$\alpha x + y \in W$$

Suppose  $B$  be the basis of vector space  $W$  and  $C$  be the basis of  $V$  such that,

$$B = (v_1, v_2), C = (x_1, x_2)$$

Then, the basis for  $V_n$  will be,

$$C_n = \left( \frac{1}{n}x_1, \frac{1}{n}x_2 \right)$$

Now, approximate the area of  $\Pi(B)$

Suppose  $[V:W] = r$

Then,

$$\begin{aligned} [V_n:W] &= [V_n:V][V:W] \\ &= n^2 r \end{aligned}$$

Since,  $W \subset V$  and  $B$  is the basis of  $W$ , then  $[V:W]$  is finite and is equal to the number of elements of  $V$  in the region  $\Pi'(B)$ .

Where,  $\Pi'(B)$  is the set of linear combinations  $r_1 v_1 + r_2 v_2$  with  $0 \leq r_i < 1$ .

Then,

$$\begin{aligned} \Delta W &\approx n^2 r \Delta(V_n) \\ &= r \Delta(V) \\ &= [V:W] \Delta(V) \end{aligned}$$

Then, from the above,

$$[V:W] = \frac{\Delta W}{\Delta(V)}$$

This implies that, the index  $[V:W]$  is the ratio of the area of the parallelogram  $\Pi(B)$  of linear combinations  $r_1 v_1 + r_2 v_2$  with  $0 \leq r_i < 1$  and the area of the set of linear combinations  $r_1 v_1 + r_2 v_2$  with  $0 \leq r_i < 1$ .

Hence, **the required way to determine the index  $[V:W]$  is that find the ratio of the area of the parallelogram  $\Pi(B)$  of linear combinations  $r_1 v_1 + r_2 v_2$  with  $0 \leq r_i < 1$  to the area of the set of linear combinations  $r_1 v_1 + r_2 v_2$  with  $0 \leq r_i < 1$ .**

10. a

Consider the provided statement to prove that the provided ring is a finite-dimensional vector space over  $\mathbb{C}$ .

As it is also provided that  $f$  and  $g$  is polynomials function over  $\mathbb{C}[x, y]$  and it has no common factor.

[Comment](#)

### Step 2 of 2 ^

Now as  $f$  and  $g$  has no common factor therefore greatest common divisor of  $(f, g)$  is 1. By using Bezout's identity, there is an integer  $h_1(x, y)$  and  $h_2(x, y)$  such that,

$$h_1(x, y) \cdot f(x, y) + h_2(x, y) \cdot g(x, y) = 1$$

As the provided ring is,

$$R = \frac{\mathbb{C}[x, y]}{(f, g)}$$

$$\frac{\mathbb{C}[x, y]}{(f, g)} = [(f, g) + h \mid h \in [x, y]]$$

As the provided ring is quotient ring, then any multiple of  $(f, g)$  is observed by  $\langle f, g \rangle$ .

So that there are finite number of value and the remaining values are consumed by set. Therefore, finite number of vector space over  $\mathbb{C}$ . Hence, provided ring is finite dimensional vector space over  $\mathbb{C}$  is **proved**.

11. a

(a)

To prove that  $f' = 0$  if and only if  $f$  is a square,

Suppose  $\mathbb{F}_2[x]$  be the ring of the polynomials and  $f \in \mathbb{F}_2[x]$  be a square polynomial such that,

$$f(x) = (x^2 + bx + c)^2$$

Then, either  $b \equiv 0 \pmod{2}$  or  $b \equiv 1 \pmod{2}$  and  $c \equiv 0 \pmod{2}, c \equiv 1 \pmod{2}$ .

Now,

$$f'(x) = 2(x^2 + bx + c)(2x + b)$$

Since, the polynomials and  $f \in \mathbb{F}_2[x]$  be a square polynomial.

Then,

$$\begin{aligned} f'(x) &= 2(x^2 + bx + c)(2x + b) \\ &\equiv 0 \pmod{2} \end{aligned}$$

Therefore, it concludes that, when  $f \in \mathbb{F}_2[x]$  be a square polynomial then  $f' = 0$ .

Sufficient part,

Suppose  $\mathbb{F}_2[x]$  be the ring of the polynomials and the polynomial  $f \in \mathbb{F}_2[x]$  of degree two is such that,

$$f' = 0$$

That is,

$$f' \equiv 0 \pmod{2}$$

This implies that, the coefficient of leading term and the constant term is multiple of 2 in the polynomial of  $f'$ .

Since, the coefficient of leading term of  $f'$  is the degree of the polynomial  $f$  by antiderivative rule.

Then, the polynomial  $f \in \mathbb{F}_2[x]$  must be of the form,

$$f(x) = (x^2 + bx + c)^2$$

Hence, it is proved that  $f' = 0$  if and only if  $f$  is a square in ring of the polynomials  $\mathbb{F}_2[x]$ .

(b)

To prove that either  $\gcd(f, g)$  or  $\gcd(f, g-1)$  is a proper factor of the polynomial  $f$ ,

Suppose  $\mathbb{F}_2[x]$  be the ring of the polynomials and  $f \in \mathbb{F}_2[x]$  of degree  $n$ .

Since, there is a polynomial  $g$  of the degree at most  $n$  such that,

$$g^2 - g \equiv 0 \pmod{f}$$

And  $f = uv$  such that  $u$  and  $v$  are relatively prime to each other.

Then,

$$\gcd(u, v) = 1$$

Then, from the congruence relation,

$$g^2 - g \equiv 0 \pmod{f}$$

This implies that,

$$f \mid g^2 - g$$

This implies that,

$$f \mid g(g-1)$$

Then, either  $f \mid g$  or  $f \mid (g-1)$

This implies that either  $g$  is multiple of  $f$  or  $(g-1)$  is multiple of  $f$ .

That is, either  $f$  is a factor of  $g$  or  $f$  is factor of  $(g-1)$ .

This implies that either  $\gcd(f, g) \neq 1$  or  $\gcd(f, g-1) \neq 1$ .

This implies that, either  $\gcd(f, g)$  or  $\gcd(f, g-1)$  is a proper factor of  $f$ .

Hence, it is proved that either  $\gcd(f, g)$  or  $\gcd(f, g-1)$  is a proper factor of the polynomial  $f$ .

(c)

To factor the polynomial  $(x^9 + x^6 + x^4 + 1)$  in the polynomial ring  $\mathbb{F}_2[x]$ ,

Suppose  $\mathbb{F}_2[x]$  be the ring of the polynomials and  $f \in \mathbb{F}_2[x]$  such that,

$$f(x) = x^9 + x^6 + x^4 + 1$$

Then, by Berlekamp algorithm, for any monic integer polynomial  $f$  whose residue modulo  $p$  is the product of relative prime monic polynomials in  $\mathbb{F}_p[x]$ , there will be a unique way to factor  $f$  modulo any power of  $p$ .

That is, factor the polynomial  $f(x)$  in modulo 2.

Now, use MAPLE to factor the polynomial  $f(x)$  in modulo 2.

$$\text{Factor}(x^9 + x^6 + x^4 + 1) \bmod(2) \\ (x^8 + x^7 + x^6 + x^3 + x^2 + x + 1)(x + 1)$$

Therefore, it concludes that the factorization of the polynomial in  $\mathbb{F}_2[x]$  is,

$$(x^9 + x^6 + x^4 + 1) \equiv (x^8 + x^7 + x^6 + x^3 + x^2 + x + 1)(x + 1) \pmod{2}$$

Hence, the required factorization of the polynomial in the polynomial ring  $\mathbb{F}_2[x]$  is

$$\boxed{(x^9 + x^6 + x^4 + 1) = (x^8 + x^7 + x^6 + x^3 + x^2 + x + 1)(x + 1)}.$$

## Chapter 13

### Section 1

1. a

Yes, because its irreducible polynomial over  $\mathbb{Q}$  is  $x^2 - x - 1$ . In order to see how we chose this polynomial, note that the roots  $x_1, x_2$  of  $ax^2 + bx + c$  are given by

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Now, the form of  $\frac{1+\sqrt{5}}{2}$ , as it has to be a root of that polynomial, immediately suggests  $a = 1$ ,  $b = -1$ , and since we want  $b^2 - 4ac = 5$  then also  $c = -1$ .

#### Step 2

2 of 3

**Remark.** This result also follows from applying **Proposition 13.1.6** with  $\delta = \sqrt{5}$ , as  $5 \equiv 1 \pmod{4}$ .

#### Result

3 of 3

Yes, since it is a root of  $x^2 - x - 1$ . Click for more details.

2. a

As per **Proposition 13.1.6**, we have two separate cases: when  $d \equiv 2 \pmod{4}$  or  $d \equiv 3 \pmod{4}$ , and when  $d \equiv 1 \pmod{4}$ .

## Step 2

2 of 4

First let  $d \equiv 2 \pmod{4}$  or  $d \equiv 3 \pmod{4}$ , then the set of algebraic integers of  $\mathbb{Q}[\sqrt{d}]$  is given by

$$R = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}.$$

We first have to show that  $R$  is closed under addition and multiplication. Closure under addition follows from the equality

$$(a + b\sqrt{d}) + (c + e\sqrt{d}) = (a + c) + (b + e)\sqrt{d},$$

and noting that  $\mathbb{Z}$  is also closed under addition. Closure under multiplication also follows from the equality

$$(a + b\sqrt{d})(c + e\sqrt{d}) = ac + ae\sqrt{d} + bc\sqrt{d} + bed = (ac + bed) + (ae + bc)\sqrt{d},$$

and noting that  $\mathbb{Z}$  is also closed under addition and multiplication.

We continue to verify that  $(R, +)$  is an abelian group. We have  $0 \in R$  (for the choice of  $a = b = 0$ ), if  $a + b\sqrt{d} \in R$  then  $-a - b\sqrt{d} \in R$  and hence  $R$  contains additive inverses, and furthermore addition in  $R$  is commutative and associative because addition of complex numbers is commutative and associative.

Finally, let us show that  $(R \setminus \{0\}, \cdot)$  is a monoid (i.e.  $\cdot$  is associative, and there is a neutral element) and that  $\cdot$  distributes over  $+$ . The associativity and distributivity follow from corresponding properties of complex numbers, while  $a + b\sqrt{d}$  is the neutral element for multiplication for the choice  $b = 0$  and  $a = 1$ .

Now let  $d \equiv 1 \pmod{4}$ , then the set of algebraic integers of  $\mathbb{Q}[\sqrt{d}]$  is given by

$$R = \left\{ a + b \left( \frac{1 + \sqrt{d}}{2} \right) : a, b \in \mathbb{Z} \right\}.$$

Note that if we prove that  $R$  is closed under addition and multiplication everything else follows roughly as much as in the first case. We observe that the identity

$$a + b \left( \frac{1 + \sqrt{d}}{2} \right) + c + e \left( \frac{1 + \sqrt{d}}{2} \right) = a + c + (b + e) \left( \frac{1 + \sqrt{d}}{2} \right)$$

proves the closure under addition. Now, for multiplication, note that we have

$$\begin{aligned} \left( a + b \left( \frac{1 + \sqrt{d}}{2} \right) \right) \left( c + e \left( \frac{1 + \sqrt{d}}{2} \right) \right) &= ac + (ae + bc) \left( \frac{1 + \sqrt{d}}{2} \right) + be \left( \frac{1 + \sqrt{d}}{2} \right)^2 \\ &= ac + (ae + bc) \left( \frac{1 + \sqrt{d}}{2} \right) + be \left( \frac{1 + 2\sqrt{d} + d^2}{4} \right) \\ &= ac + (ae + bc) \left( \frac{1 + \sqrt{d}}{2} \right) + be \left( \frac{1 + 2\sqrt{d} + d^2 + 1 - 1}{4} \right) \\ &= ac + (ae + bc) \left( \frac{1 + \sqrt{d}}{2} \right) + be \left( \frac{2 + 2\sqrt{d} + d^2 - 1}{4} \right) \\ &= ac + (ae + bc) \left( \frac{1 + \sqrt{d}}{2} \right) + be \left( \frac{1 + \sqrt{d}}{2} \right) + be \left( \frac{d^2 - 1}{4} \right) \\ &= \left( ac + be \left( \frac{d^2 - 1}{4} \right) \right) + (ae + bc + be) \left( \frac{1 + \sqrt{d}}{2} \right) \end{aligned}$$



which proves that  $R$  is closed under multiplication after noting that, since  $d \equiv 1 \pmod{4}$ , then also  $d^2 \equiv 1 \pmod{4}$ , and hence 4 divides  $d^2 - 1$ , i.e.  $\frac{d^2-1}{4}$  is an integer. As we have stated, the other properties follow analogously as in the first case, and therefore this finishes our proof.

## Result

4 of 4

For each of the cases, when  $d \equiv 2, 3 \pmod{4}$  and when  $d \equiv 1 \pmod{4}$ , we verify the ring axioms directly.  
[Click to see more details.](#)

### 3. a

#### (a)

Let  $f(x)$  be a monic integer polynomial such that  $f(\alpha) = 0$ . If  $f$  is irreducible, then we are done; if not, then as  $\mathbb{Z}[x]$  is a unique factorization domain (this is **Theorem 12.3.8**, note that  $f(x)$  is monic and hence there are no integer primes in our product), then there are irreducible polynomials  $q_1(x), \dots, q_n(x)$  such that

$$f(x) = q_1(x) \cdots q_n(x).$$

As  $f(\alpha) = 0$  then  $q_1(\alpha) \cdots q_n(\alpha) = 0$ , which implies there is at least one  $i = 1, \dots, n$ , such that  $q_i(\alpha) = 0$ . Now, by Gauss' lemma (or rather its immediate consequence, **Proposition 12.3.7**), we have that as  $q_i(x)$  is irreducible over  $\mathbb{Q}[x]$  as well. Therefore,  $q_i(x)$  is monic integer polynomial irreducible over  $\mathbb{Q}[x]$ , and  $q_i(\alpha) = 0$ , from which we can conclude that  $\alpha$  is an algebraic integer.

#### (b)

Since  $\alpha$  is a root of  $f(x)$ , then

$$a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_1 \alpha + a_0 = 0. \quad (1)$$

Multiplying both sides of (1) by  $a_n^{n-1}$  we obtain

$$(a_n \alpha)^n + a_{n-1} (a_n \alpha)^{n-1} + \cdots + a_n^{n-2} a_1 (a_n \alpha) + a_n^{n-1} a_0 = 0,$$

showing that  $a_n \alpha$  is the root of

$$g(x) = x^n + a_{n-1} x^{n-1} + a_n a_{n-2} x^{n-2} + \cdots + a_n^{n-2} a_1 x + a_n^{n-1} a_0$$

which is a monic integer polynomial, and hence by **(a)** we obtain that  $a_n \alpha$  is an algebraic integer.

(c)

Suppose first that  $a_0 = \pm 1$ . Then, we have

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha = \mp 1,$$

where dividing both sides by  $\alpha$  we obtain

$$\alpha^{n-1} + a_{n-1}\alpha^{n-2} + \cdots + a_1 = \mp \alpha^{-1}.$$

Note that the left-hand side here is just product and sum of algebraic integers (as both integers and  $\alpha$  are algebraic integers), and since algebraic integers are a ring, then they are closed under finite products and sums; therefore since the left-hand side is an algebraic integer, then so is the right-hand side, which is what we wanted to show. (If the sign of the right-hand side is  $-$  then we again use the fact that algebraic integers are a ring, and hence are closed under additive inverses.)

Note first that this is incorrect as stated, as we require also the polynomial in question to be irreducible. For example  $i$  (the imaginary unit) is an algebraic integer (since it is the root of  $x^2 + 1$ ), and its inverse  $i^{-1} = -i$  is also an algebraic integer, but if we drop the irreducibility requirement then we can consider that monic polynomial  $x^3 + 3x^2 + x + 3 = (x^2 + 1)(x + 3)$  has  $i$  as its root but does not have the constant part equal to  $\pm 1$ .

Suppose now that  $\alpha^{-1}$  is an algebraic integer and that  $\alpha$  satisfies the hypotheses of the exercise, i.e.

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0.$$

Dividing both sides by  $\alpha^n$  we obtain

$$1 + a_{n-1}\frac{1}{\alpha} + \cdots + a_1\left(\frac{1}{\alpha}\right)^{n-1} + a_0\left(\frac{1}{\alpha}\right)^n = 0,$$

and hence

$\alpha^{-1}$  is the root of

$$g(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + 1.$$

Denote  $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ , then  $g(x) = x^n f(1/x)$ . Since (as the opening remark clarifies) we must assume that  $f(x)$  is irreducible, we want to show that this implies that  $g(x)$  is irreducible.

Suppose that  $g(x) = p(x)q(x)$  for some nonconstant polynomials  $p(x)$  and  $q(x)$ , i.e.

$$p(x)q(x) = x^n f(1/x).$$

Substituting  $y = 1/x$  we obtain

$$f(y) = y^n p(1/y)q(1/y). \quad (2)$$

We know that the  $\deg p + \deg q = \deg g = n$ , and hence we can write (2) as

$$f(y) = (y^{\deg p} p(1/y))(y^{\deg q} q(1/y)),$$

where  $y^{\deg p} p(1/y)$  and  $y^{\deg q} q(1/y)$  are nonconstant polynomials, contradicting the irreducibility of  $f(y)$ .

Therefore,  $g(x)$  is irreducible; it is monic only if its leading coefficient  $a_0$  is either 1 or  $-1$  -- in the latter case we can divide the whole polynomial through with  $-1$  without dividing its roots, while we can't divide it through anything else because the constant part of  $g(x)$  is 1, and therefore dividing with something other than  $\pm 1$  would make the constant part a non-integer.

But as  $g(x)$  must be monic since  $\alpha^{-1}$  is its root and is an algebraic integer, it follows that  $a_0 = \pm 1$ .

(c)

Suppose first that  $a_0 = \pm 1$ . Then, we have

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha = \mp 1,$$

where dividing both sides by  $\alpha$  we obtain

$$\alpha^{n-1} + a_{n-1}\alpha^{n-2} + \cdots + a_1 = \mp \alpha^{-1}.$$

Note that the left-hand side here is just product and sum of algebraic integers (as both integers and  $\alpha$  are algebraic integers), and since algebraic integers are a ring, then they are closed under finite products and sums; therefore since the left-hand side is an algebraic integer, then so is the right-hand side, which is what we wanted to show. (If the sign of the right-hand side is  $-$  then we again use the fact that algebraic integers are a ring, and hence are closed under additive inverses.)

Now we deal with the other direction. Note first that this is incorrect as stated, as we require also the polynomial in question to be irreducible. For example  $i$  (the imaginary unit) is an algebraic integer (since it is the root of  $x^2 + 1$ ), and its inverse  $i^{-1} = -i$  is also an algebraic integer, but if we drop the irreducibility requirement then we can consider that monic polynomial  $x^3 + 3x^2 + x + 3 = (x^2 + 1)(x + 3)$  has  $i$  as its root but does not have the constant part equal to  $\pm 1$ .

Suppose now that  $\alpha^{-1}$  is an algebraic integer and that  $\alpha$  satisfies the hypotheses of the exercise, i.e.

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0.$$

Dividing both sides by  $\alpha^n$  we obtain

$$1 + a_{n-1}\frac{1}{\alpha} + \cdots + a_1\left(\frac{1}{\alpha}\right)^{n-1} + a_0\left(\frac{1}{\alpha}\right)^n = 0,$$

and hence

$\alpha^{-1}$  is the root of

$$g(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + 1.$$

Denote  $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ , then  $g(x) = x^n f(1/x)$ . Since (as the opening remark clarifies) we must assume that  $f(x)$  is irreducible, we want to show that this implies that  $g(x)$  is irreducible.

Suppose that  $g(x) = p(x)q(x)$  for some nonconstant polynomials  $p(x)$  and  $q(x)$ , i.e.

$$p(x)q(x) = x^n f(1/x).$$

Substituting  $y = 1/x$  we obtain

$$f(y) = y^n p(1/y)q(1/y). \quad (2)$$

We know that the  $\deg p + \deg q = \deg g = n$ , and hence we can write (2) as

$$f(y) = (y^{\deg p} p(1/y))(y^{\deg q} q(1/y)),$$

where  $y^{\deg p} p(1/y)$  and  $y^{\deg q} q(1/y)$  are nonconstant polynomials, contradicting the irreducibility of  $f(y)$ .

Therefore,  $g(x)$  is irreducible; it is monic only if its leading coefficient  $a_0$  is either 1 or  $-1$  -- in the latter case we can divide the whole polynomial through with  $-1$  without dividing its roots, while we can't divide it through anything else because the constant part of  $g(x)$  is 1, and therefore dividing with something other than  $\pm 1$  would make the constant part a non-integer.

But as  $g(x)$  must be monic since  $\alpha^{-1}$  is its root and is an algebraic integer, it follows that  $a_0 \pm 1$ .

## Result

6 of 6

In **(a)** part we use the unique factorization of polynomials and Gauss' lemma, in **(b)** we use **(a)** part and some algebraic manipulations, while in **(c)** we use the fact that algebraic integers form a ring for one direction, and irreducibility of  $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$  implying the irreducibility of  $g(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + 1$  in the other. Click to see more details.

## 4. a

Let  $d$  and  $d'$  be two distinct squarefree integers. We want to show that  $\mathbb{Q}[\sqrt{d}]$  and  $\mathbb{Q}[\sqrt{d'}]$  are distinct -- it is sufficient to show that  $\sqrt{d} \notin \mathbb{Q}[\sqrt{d'}]$ . Suppose on the contrary, that  $\sqrt{d} \in \mathbb{Q}[\sqrt{d'}]$ , then there are rational numbers  $a$  and  $b$  such that

$$a + b\sqrt{d'} = \sqrt{d}, \quad (1)$$

where it is not hard to see that we must have both  $a \neq 0$  and  $b \neq 0$ . First one follows from the fact that if  $a = 0$  then by squaring both sides it follows that

$$b^2 d' = d,$$

but since  $d$  is squarefree then  $b^2 = 1$  and therefore  $d' = d$ , contradicting our assumption. Similarly, if  $b = 0$  that implies that  $\sqrt{d}$  is a rational number, again a contradiction. Therefore, since  $ab \neq 0$ , by squaring (1) we obtain

$$a^2 + 2ab\sqrt{d'} + b^2 d' = d,$$

which after some straightforward algebraic transformations turns into

$$\sqrt{d'} = \frac{d - a^2 - b^2 d'}{2ab}.$$

But the left-hand side here is irrational and the right-hand side is rational, i.e. a contradiction with existence of such  $a$  and  $b$ , showing that  $\mathbb{Q}[\sqrt{d}] \neq \mathbb{Q}[\sqrt{d'}]$ . (In fact, more can be shown about this, not only that they're distinct, but also that they're not isomorphic.)

Therefore, since as remarked in the chapter we have that  $\mathbb{Q}[\sqrt{d}] = \mathbb{Q}[\sqrt{m^2 d}]$  for any  $m \in \mathbb{Z}$ , we see that  $\mathbb{Q}[\sqrt{d}]$  and  $\mathbb{Q}[\sqrt{d'}]$  are distinct if and only if squarefree parts of  $d$  and  $d'$  -- i.e. what remains of them after we divide both by their own square factors -- are distinct.

## Result

2 of 2

We show that  $\mathbb{Q}[\sqrt{d}]$  and  $\mathbb{Q}[\sqrt{d'}]$  are distinct if and only if squarefree parts of  $d$  and  $d'$  are distinct. Click for more details.



## Section 2

### 1. a

Recall that the norm function in  $R = \mathbb{Z}[\sqrt{-5}]$  is of the form

$$N(a + b\sqrt{-5}) = a^2 + 5b^2,$$

where  $a$  and  $b$  are integers. Therefore, we have

$$N(2) = 4, N(3) = 9, \text{ and } N(1 \pm \sqrt{-5}) = 5.$$

First note that if an element  $\alpha \in R$  is not irreducible then there are non-unit elements  $\beta$  and  $\gamma$  such that  $\alpha = \beta\gamma$ , but then

$$N(\alpha) = N(\beta)N(\gamma),$$

so that, as  $R$  is an imaginary quadratic ring, and  $\beta$  and  $\gamma$  are non-units, we have that  $N(\beta)$  is a positive integers greater than 1 dividing  $N(\alpha)$ , and similarly for  $N(\gamma)$ . This immediately shows that  $1 \pm \sqrt{-5}$  is irreducible, as its norm is a prime number, which implies that in any decomposition of the form  $1 \pm \sqrt{-5} = \beta\gamma$  we must have  $\beta$  or  $\gamma$  be a unit.

Now we investigate the cases of 2 and 3. Note that if  $2 = \beta\gamma$  is a decomposition into nonunits (in  $R$ ), then by apply norm function to both sides we see immediately that we must have that

$$N(\beta) = N(\gamma) = 2,$$

and similarly if  $3 = \delta\tau$  then we see that

$$N(\delta) = N(\tau) = 3.$$

Therefore, irreducibility of 2 and 3 would follows if we were to prove that equations

$$a^2 + 5b^2 = 2 \text{ and } a^2 + 5b^2 = 3 \tag{1}$$

have no solutions with  $a, b$  integers. This is, however, straightforward from considering that  $a^2 > 0$  for  $a \in \mathbb{Z}$  and  $5b^2 \geq 5$  for  $b \neq 0$ , which implies that in both of equations (1) we must have  $b = 0$ , i.e.

$$a^2 = 2 \text{ and } a^2 = 3.$$

But neither 2 nor 3 are squares in  $\mathbb{Z}$ , showing that neither of those equations have any solutions in  $\mathbb{Z}$ , and therefore showing 2 and 3 are units in  $R$ .

### Step 2

2 of 3

Now suppose  $\alpha \in R$  is a unit, then  $N(\alpha) = 1$ , so that if  $\alpha = a + b\sqrt{-5}$ , then

$$a^2 + 5b^2 = 1.$$

By completely analogues reasoning as in the previous paragraph we see that the only integer solutions to this are  $a = 1, b = 0$ , and  $a = -1, b = 0$ .

### Result

3 of 3

We show the irreducibility of 2, 3 and  $1 \pm \sqrt{-5}$  by considering their norms and showing how their decomposition as integers leads one to conclude that those numbers are irreducible in  $R$ . We do the same for showing only units in  $R$  are  $\pm 1$ , which then follows from the fact that  $a^2 + 5b^2 = 1$  has two solutions with  $a, b$  integers, namely  $a = 1, b = 0$  and  $a = -1, b = 0$ . Click to see more details.

### 2. a

As suggested by the theorem classifying which imaginary quadratic rings are unique factorization domains (**Theorem 13.2.5**), we are to prove the following proposition:

Let  $R$  be the ring of integers in an imaginary quadratic number field  $\mathbb{Q}[\sqrt{d}]$ , where  $d \equiv 2 \pmod{4}$ . Then  $R$  is not a unique factorization domain for any  $d < -2$ .

## Step 2

2 of 3

**Proof.** Note first that as  $d \equiv 2 \pmod{4}$ , we have that the elements of  $R$  are of the form  $a + b\sqrt{-d}$  with  $a, b$  integers. Furthermore, suppose that  $\eta = m + n\sqrt{-d}$  is a unit of  $R$ , then

$$m^2 + dn^2 = 1,$$

implying that  $m^2 = 1$  (since  $d > 1$  and therefore  $dn^2 > 1$  except for  $n = 0$ ), so that we see that the units of  $R$  are  $\pm 1$ .

Let  $e = \frac{4-d}{2}$ , which is an integer since  $d \equiv 2 \pmod{4}$  implies that  $d$  is even. Then

$$2e = 4 - d = (2 - \sqrt{d})(2 + \sqrt{d}).$$

Hence  $4 - d$  has two factorizations in  $R$ , and since  $d < -2$  then there is no element of  $R$  whose norm is equal to 2. This implies that 2 is an irreducible element of  $R$ , so that if  $R$  was a unique factorization domain this would mean that 2 divides either  $2 - \sqrt{d}$  or  $2 + \sqrt{d}$  in  $R$ . But as  $\frac{1}{2}(2 - \sqrt{d}) = 1 + \frac{\sqrt{d}}{2}$  is not in  $R$  when  $d \equiv 2 \pmod{4}$ , this cannot be, which concludes our proof.

## Result

3 of 3

We follow the similar reasoning as in the cases when  $d \equiv 3 \pmod{4}$ , which was proved in the text, but considering  $e = (4 - d)/2$  instead of  $e = (1 - d)/2$ . Click for more details.

# Section 3

## 1. a

Lattice is the fundamental algebraic structures which consists of a partially ordered set in which every two elements have a unique supremum and a unique infimum

[Comment](#)

## Step 2 of 3

The lattice basis  $(\alpha + \alpha\delta)$  of the principal ideal  $(\alpha)$  is obtained from the lattice basis  $(1, \delta)$  of the unit ideal  $R$  by multiplying  $\alpha$ .

Now, write  $\alpha$  in polar coordinates;

$$\alpha = re^{i\theta}$$

Then multiplication by  $\alpha$  rotates the complex plane through the angle  $\theta$  and stretches by the factor  $r$ .

So, all principal ideals are similar geometric figures.



Also, the lattice with basis;

$$\left( \alpha, \frac{1}{2}(\alpha + \alpha\delta) \right)$$

This is obtained from the matrix  $(2, (1+\delta))$  by multiplying with  $\frac{1}{2}\alpha$

The ideal  $(2, (1+\delta))$  in the Ring  $\mathbb{Z}[\sqrt{-5}]$

Similarity classes of ideals are called ideal classes, and the number of ideal classes is the class number of Ring  $R$ .

Hence,  $(\alpha, \delta)$  are ideal for the lattice  $(2, (1+\delta))$ .

2. a

We check for, each of these, whether they're closed under multiplication and addition.

## Step 2

2 of 5

(a)

Suppose  $I = (5, 1 + \sqrt{-5})$  were an ideal. As it is closed under multiplication, we have that  $\sqrt{-5}(1 + \sqrt{-5}) \in I$ . Therefore

$$\sqrt{-5}(1 + \sqrt{-5}) = -5 + \sqrt{-5} = 5n + (1 + \sqrt{-5})m,$$

for some integers  $n$  and  $m$ . This yields two equations,

$$\begin{aligned} -5 &= 5n + m \\ 1 &= m, \end{aligned}$$

i.e. we have  $m = 1$  and we need to obtain  $n$  from  $5n + 1 = -5$ . But the only solutions to this equation is  $n = \frac{-6}{5}$ , which is not an integer. Therefore,  $I$  is not closed under multiplication and hence not an ideal.

(b)

Suppose  $I = (7, 1 + \sqrt{-5})$  was an ideal. Then analogously as in (a) (since the second generating element is the same) we would obtain two equations

$$\begin{aligned} -5 &= 7n + m \\ 1 &= m, \end{aligned}$$

where the only solution is  $m = 1$  and  $n = \frac{-6}{7}$ , where  $n \notin \mathbb{Z}$ .

(c)

Let  $I = (4 - 2\sqrt{-5}, 2 + 2\sqrt{-5}, 6 + 4\sqrt{-5})$ . We want to show that for  $i, j \in I$ , and for  $a + b\sqrt{-5}$  and  $c + d\sqrt{-5}$  with  $a, b, c, d$  integers, we have

$$(a + b\sqrt{-5})i + (c + d\sqrt{-5})j \in I.$$

Note that

$$(a + b\sqrt{-5})i + (c + d\sqrt{-5})j = ai + cj + (b + d)\sqrt{-5},$$

where  $ai, cj$  and  $b + d$  are in  $I$  because  $I$  is a lattice, so that it is sufficient to show that for any  $k \in I$  we have  $k\sqrt{-5} \in I$ . In particular, since any  $k \in I$  can be written as an integer combination of the generators of  $I$ , it is sufficient to show that for each generator  $i$ , we have  $i\sqrt{-5} \in I$ . We check this case by case.

First, note that an integer combination of the generators is of the form

$$a(4 - 2\sqrt{-5}) + b(2 + 2\sqrt{-5}) + c(6 + 4\sqrt{-5}) = (4a + 2b + 6c) + (-2a + 2b + 4c)\sqrt{-5}.$$

Now we compute

$$\sqrt{-5}(4 - 2\sqrt{-5}) = 10 + 4\sqrt{-5}, \quad (1)$$

$$\sqrt{-5}(2 + 2\sqrt{-5}) = -10 + 2\sqrt{-5}, \quad (2)$$

$$\sqrt{-5}(6 + 4\sqrt{-5}) = -20 + 6\sqrt{-5}, \quad (3)$$

which translate into three pairs of equations; for (1) we have

$$4a + 2b + 6c = 10 \text{ and } -2a + 2b + 4c = 4,$$

where we can notice a solutions  $a = 1, b = 3$ . For (2) we obtain

$$4a + 2b + 6c = -10 \text{ and } -2a + 2b + 4c = 2,$$

where we notice a solution  $a = -2, b = -1$ ; and lastly, for (3) we have

$$4a + 2b + 6c = -20 \text{ and } -2a + 2b + 4c = 6,$$

where a solution is given by  $a = -4, b = 1, c = -1$ .

## Result

5 of 5

In (a) and (b) we show that they are not ideals, while in (c) we show that it is an ideal. Click to see more details.

3. a

**Given:**  $A$  is an ideal of the ring of integers  $R$  in an imaginary quadratic field.

**To Prove:** There is a lattice basis for  $A$ , one of whose elements is an ordinary positive integer.

**Proof:** Let us assume that  $x \in A$ .

Then note that  $\bar{x}x \in A$  is a positive integer.

This follows that

$$\bar{x}x \in A \cap \mathbb{N}.$$

Therefore we have the set  $A \cap \mathbb{N}$  is non-empty.

Then by **Well ordering property** let us choose  $a$  in  $A \cap \mathbb{N}$  such that  $a$  is the minimal element of  $A \cap \mathbb{N}$ .

Let us now choose  $y \in A$  such that  $y$  is  $\mathbb{Z}$ -linearly independent from  $x$ .

Now consider the parallelogram

$$P(a, y) := \{sa + ty \mid 0 \leq s, t \leq 1\}.$$

Now notice that  $A \cap P(a, y)$  is a finite set. Therefore we can choose  $b$  in  $A \cap P(a, y)$  with minimal positive imaginary part.

We will propose to prove that  $a$  and  $b$  form a lattice basis for  $A$ .

Now notice that

$$A \cap P(a, b) = \{0, a, b, a + b\}.$$

If possible let us assume  $c \in A \cap P(a, b) - \{0, a, b, a + b\}$ .

**Case-1:**  $c \in \mathbb{Z}$ .

But we have  $c < a$  and violating minimality of  $a$ .

So in this case  $c$  is null.

**Case-2:**  $0 < \text{Im}(a) < \text{Im}(b)$ .

This also violating minimality of  $b$ .

Hence in this case also  $b$  is null.

Therefore we can conclude that

$$A \cap P(a, b) = \{0, a, b, a + b\}.$$

## Step 2

2 of 3

Now consider the parallelogram  $P(a, b)$  by  $\mathbb{Z}$ -linear combinations of  $a, b$ .

And none of these parallelograms contain anything in  $A$  other than linear combinations of  $a$  and  $b$  for otherwise we can move that point into the parallelogram  $P(a, b)$  with an appropriate  $\mathbb{Z}$ -linear combination of  $a, b$ .

Thus,

$a, b$  form a lattice basis for  $A$ .

Therefore we proved that there is a lattice basis for  $A$ , one of whose elements is an ordinary positive integer.

This completes the proof.

## Result

Considering a parallelogram  $P(a, b)$  we have shown that  $a, b$  form a lattice basis for  $A$ .

## 4. a

Let  $R$  be a ring then the subring is defined as the subset of a ring that is itself a ring when binary operations of addition and multiplication on  $R$  are restricted to the subset which shares the same multiplicative identity as in  $R$ .

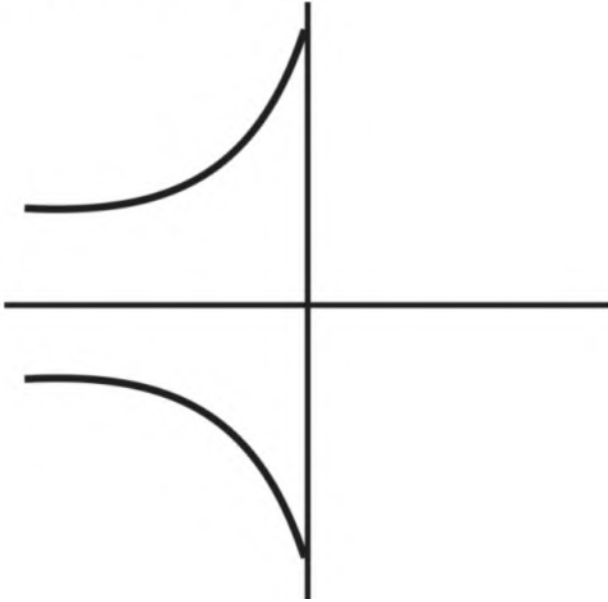
a.

Consider the ring;

$$R = \mathbb{Z}\sqrt{-3}$$

$$\text{Let, } \mathbb{Z}\sqrt{-3} = a + b\sqrt{-3}$$

The  $\mathbb{Z}\sqrt{-3}$  Lattice is drawn below:



Since, the norm of an algebraic integer is an ordinary integer, a unit must have norm  $N(\alpha) = \pm 1$

However, it can be represented in  $R$  as a lattice in  $R^2$  by associating to the algebraic integer  $a + b\sqrt{d}$  the point  $(u, v)$  of  $R^2$ , where  $u = a + b\sqrt{d}$  and  $v = a - b\sqrt{d}$ .

Thus  $\alpha = 1 + \sqrt{-3}$  is the ideal unit in the ring  $R = \mathbb{Z}\sqrt{-3}$ , is the point on the lattice that lie on one of the two hyperbolas  $uv = 1$  and  $uv = -1$ .

b.

Consider the ring;

$$R = \mathbb{Z}\left(\frac{1}{2}(1 + \sqrt{-3})\right)$$

The lattice basis  $(\alpha + \alpha\delta)$  of the principal ideal  $(\alpha)$  is obtained from the lattice basis  $(1, \delta)$  of the unit ideal  $R$  by multiplying  $\alpha$ .

If  $\alpha$  is written in polar coordinates;

$$\alpha = re^{i\theta}$$

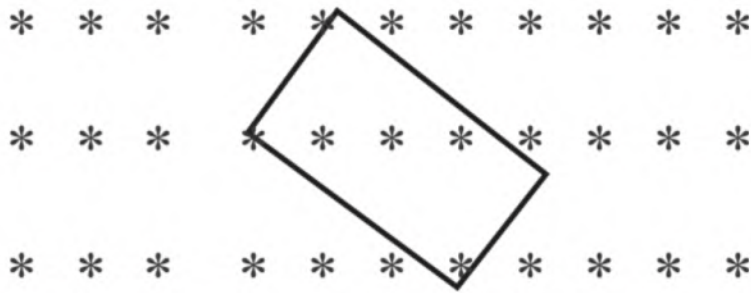
Then multiplication by  $\alpha$  rotates the complex plane through the angle  $\theta$  and stretches by the factor  $r$ .

So, all principal ideals are similar geometric figures.

Also, the lattice with basis  $\left(\alpha, \frac{1}{2}(\alpha + \alpha\delta)\right)$  is obtained from the lattice  $(3, (1 + \delta))$  by

multiplying with  $\frac{1}{2}\alpha$ .

The ideal  $(3, (1 + \delta))$  in the Ring  $\mathbb{Z} = \sqrt{-3}$  is drawn below;



Thus  $\alpha(3, (1+\delta))$  is the ideal unit in the ring  $R = \mathbb{Z}\left(\sqrt{\frac{1}{2}}(1+\sqrt{-3})\right)$ , is the point on the lattice that lie on rectangle.

c.

Consider the ring;

$$R = \mathbb{Z}\sqrt{-6}$$

Considering,  $\delta^2 = (-6)$

As per the principal of ideals,

$$2 \times 3 = 6$$

$$= (1+\delta)(1-\delta)$$

Now, by dividing  $(1+\delta)$  and  $(1-\delta)$  with 2, four ideal by factoring 6 is obtained,

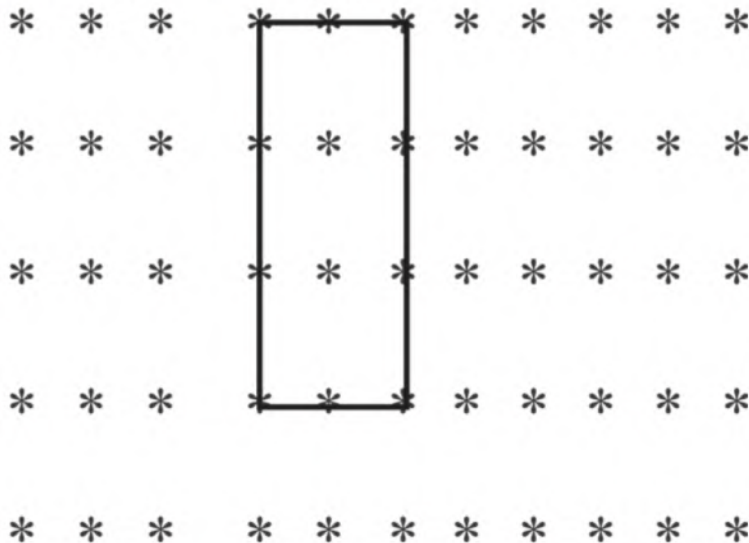
$$A = 2, (1+\delta)$$

$$A' = 2, (1-\delta)$$

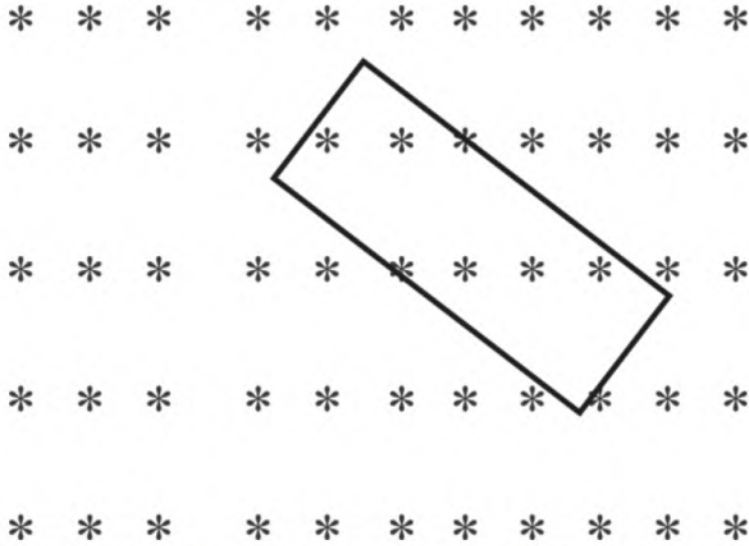
$$B = 3, (1+\delta)$$

$$B' = 3, (1-\delta)$$

The ideal  $(2, (1+\delta))$  in the Ring  $\mathbb{Z} = \sqrt{-6}$  is drawn below;



And, the ideal  $(3, (1 + \delta))$  in the Ring  $Z = \sqrt{-6}$  is drawn below;



Hence, forms rectangular lattice in the ring  $R = Z\sqrt{-6}$  for the ideal conditions  $(2, (1 + \delta))$  and  $(3, (1 + \delta))$ .

d.

Consider the ring;

$$R = Z\left(\frac{1}{2}(1 + \sqrt{-7})\right)$$

The lattice form for  $\delta = (-7)$  is as below, the lattice  $\delta = (-7)$ ;



Since at  $\delta = (-7)$ , the factors cannot be form in the ideal case of  $(2, 1 + \delta)$  or  $(3, 1 + \delta)$  so it cannot be consider as ideal in the Ring;

$$R = Z\left(\frac{1}{2}(1 + \sqrt{-7})\right) \quad R = Z\left(\frac{1}{2}(1 + \sqrt{-7})\right).$$



e.

Consider the ring;

$$R = \mathbb{Z}[\sqrt{-10}]$$

Considering the  $\delta = (-10)$ ,

As per the principal of ideals,

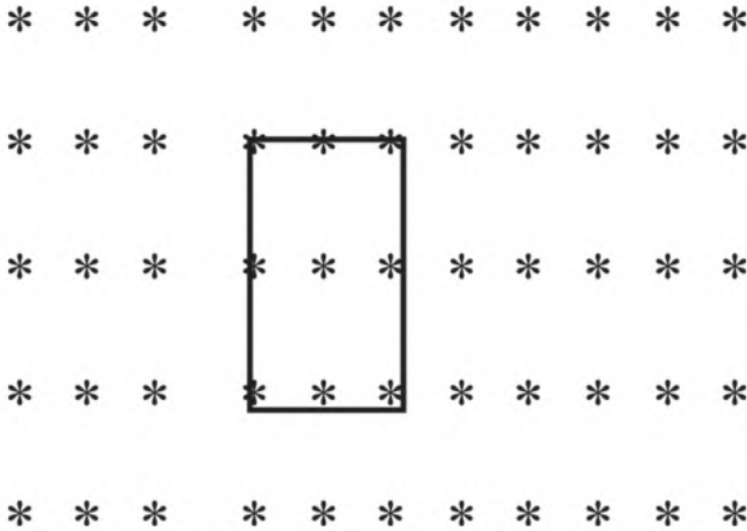
$$\begin{aligned} 2 \times 5 &= 10 \\ &= (1 + \delta)(1 - \delta) \end{aligned}$$

Now, two ideal after factorizing 10 is,

$$A = 2, 1 + \delta \text{ and } A' = 2, 1 - \delta$$

Thus the lattice in the ring  $R = \mathbb{Z}[\sqrt{-10}]$  is as follows,

The ideal  $(2, (1 + \delta))$  in the Ring  $\mathbb{Z}[\sqrt{-10}]$



Hence the  $\delta = (-10)$  is ideal in the ring  $R = \mathbb{Z}[\sqrt{-10}]$  only for  $(2, (1 + \delta))$ .

## Section 4

1. a

Consider the provided statement to find a lattice basis for the product ideal  $A, B$ .

As it is provided that  $R = \mathbb{Z}[\sqrt{-6}]$ ,  $A = (2, \delta)$  and  $B = (3, \delta)$ .

As it is known that,

$$\begin{aligned} 6 &= (2)(3) \\ &= (\sqrt{-6})(\sqrt{-6}) \end{aligned}$$

As  $\bar{A} = A$  and  $\bar{B} = B$  since  $-\sqrt{-6} \in (\sqrt{-6})$

Now it is to be shown that  $A$  and  $B$  are prime ideals. It is the way to prove that  $P$  is a prime ideal is to prove  $\frac{R}{P}$  is a field. Therefore, in this case

$$\frac{R}{A} \cong \mathbb{F}_2$$

$$\frac{R}{B} \cong \mathbb{F}_3$$

There is another way to prove that norms of element in this ideal are prime and therefore elements will be irreducible. Therefore by linear identities,

$$A^2 = (2)$$

$$B^2 = (3)$$

$$AB = (\sqrt{-6})$$

Hence,  $A^2 B^2 = (6)$

2. a

Let  $R = \mathbb{Z}[\sqrt{-5}]$ , where  $\mathbb{Z}[\sqrt{-5}] = \{a + b\delta \mid a, b \in \mathbb{Z}\}$  and  $\delta = \sqrt{-5}$ .

Then a generating set is said to form a basis of a space if it is linearly independent.

Let  $\eta = \{a_1, a_2, a_3, \dots, a_k\}$  be a set then it is linearly independent if following holds

$$\sum_{i=1}^k b_i a_i = 0$$

$$\Rightarrow b_i = 0, 1 \leq i \leq k$$

(a)

Let  $A$  be the ideal generated by  $3 + 5\delta$  and  $2 + 2\delta$ .

Then,

$$A = (3 + 5\delta, 2 + 2\delta).$$

Consider  $\alpha(3 + 5\delta) + \beta(2 + 2\delta) = 0$ , where  $\alpha, \beta \in \mathbb{Z}[\sqrt{-5}]$ .

So let,

$$\alpha = a + b\delta$$

$$\beta = c + d\delta$$

where  $a, b, c, d \in \mathbb{Z}$

Now,

$$(a + b\delta)(3 + 5\delta) + (c + d\delta)(2 + 2\delta)$$

$$= (3a + (3b + 5a)\delta - 25b) + (2c + (2d + 2c)\delta - 10d)$$

$$= (3a + 2c - 25b - 10d) + (3b + 5a + 2d + 2c)\delta$$

Since  $\alpha(3 + 5\delta) + \beta(2 + 2\delta) = 0$

So comparing terms on both sides of the equation,

$$3a - 25b + 2c - 10d = 0$$

$$5a + 3b + 2c + 2d = 0$$

Rewrite this system of equation in form of matrices.

$$\begin{pmatrix} 3 & -25 & 2 & -10 \\ 5 & 003 & 2 & 002 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Now since the reduced row echelon form of this matrix equation is given by

$$\begin{pmatrix} 1 & 0 & 28/67 & 10/67 \\ 0 & 1 & -2/67 & 28/67 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence there are no integral solution of these equations but  $a = 0, b = 0, c = 0, d = 0$ .

Since,

$$\alpha = a + b\delta$$

$$\beta = c + d\delta$$

Thus  $\alpha = 0$  and  $\beta = 0$ .

Hence linear independence holds.

Thus  $\{3 + 5\delta, 2 + 2\delta\}$  form a lattice basis for  $A$ .

Now  $\bar{A}$  is given by

$$\bar{A} = (3 - 5\delta, 2 - 2\delta)$$

So,

$$\begin{aligned} \bar{A}A &= (3 - 5\delta, 2 - 2\delta)(3 + 5\delta, 2 + 2\delta) \\ &= (134, 56 - 4\delta, 56 + 4\delta, 24) \end{aligned}$$

(b)

Let  $A$  be the ideal generated by  $4 + \delta$  and  $1 + 2\delta$ .

Then,

$$A = (4 + \delta, 1 + 2\delta).$$

Let  $\{4 + \delta, 1 + 2\delta\}$  be a linearly dependent set then there exists some  $\alpha \in \mathbb{Z}[\sqrt{-5}]$  such that

$$4 + \delta = \alpha(1 + 2\delta)$$

Now,

$$\text{Since } \alpha \in \mathbb{Z}[\sqrt{-5}]$$

So,

$$\alpha = a + b\delta \text{ for some } a, b \in \mathbb{Z}$$

Then,

$$\begin{aligned} 4 + \delta &= \alpha(1 + 2\delta) \\ &= (a + b\delta)(1 + 2\delta) \\ &= (a - 10b + (2a + b)\delta) \end{aligned}$$

Now separating the terms and comparing them yields following equation,

$$a - 10b = 4$$

$$2a + b = 1$$

On solving these equations,

$$b = -\frac{1}{3}$$

$$a = \frac{2}{3}$$

This is a contradiction since  $a, b \in \mathbb{Z}$ .

Hence  $\{4 + \delta, 1 + 2\delta\}$  is a linearly independent set.

Hence it forms a lattice basis of the given ideal  $A$ .

Now  $\bar{A}$  is given by

$$\bar{A} = (4 - \delta, 1 - 2\delta)$$

So,

$$\begin{aligned}
\bar{A}A &= (4 - \delta, 1 - 2\delta)(4 + \delta, 1 + 2\delta) \\
&= (21, 14 + 7\delta, 14 - 7\delta, 21) \\
&= (21, 14 + 7\delta, 14 - 7\delta)
\end{aligned}$$

Finally summarizing the solution,

$\{3 + 5\delta, 2 + 2\delta\}$  and  $\{4 + \delta, 1 + 2\delta\}$  forms lattice basis for  $A$ .

When  $A = (3 + 5\delta, 2 + 2\delta)$ , then  $\bar{A}A = (134, 56 - 4\delta, 56 + 4\delta, 24)$  and when  $A = (4 + \delta, 1 + 2\delta)$ , then  $\bar{A}A = (21, 14 + 7\delta, 14 - 7\delta)$ .

3. a

**Given:**  $R$  is the ring given as  $\mathbb{Z}[\delta]$ , where  $\delta = \sqrt{-5}$  and given ideals  $A$  and  $B$  as

$$A = \left(\alpha, \frac{1}{2}(\alpha + \alpha\delta)\right) \quad \text{and} \quad B = \left(\beta, \frac{1}{2}(\beta + \beta\delta)\right).$$

**To Prove:**  $AB$  is a principal ideal by finding a generator.

**Proof:** Let us recall the proposition which states that:

The algebraic integers in the quadratic field  $\mathbb{Q}[\delta]$ , with  $\delta^2 = d$  and  $d$  is square free, have the form  $\alpha = a + b\delta$ , where

- 1) If  $d \equiv 2$  or  $3$  modulo  $4$ , then  $a$  and  $b$  are integers.
- 2) If  $d \equiv 1$  modulo  $4$ , then  $a$  and  $b$  are either both integers, or both half integers.

From the aforementioned proposition it follows that

$$2 \text{ divides } \alpha \text{ and } 2 \text{ divides } \beta.$$

Again recall that if  $A$  and  $B$  are ideals of a ring  $R$  with  $A = (a)$  is a principal ideal and  $B$  is arbitrary then  $AB$  is the set of products  $ab$  with  $b \in B$ , that is

$$AB = aB.$$

Therefore we have

$$A = \frac{\alpha}{2}(2, 1 + \delta) \quad \text{and} \quad B = \frac{\beta}{2}(2, 1 + \delta).$$

Now note that if  $A$  and  $B$  are ideals of a ring  $R$  with  $\{a_1, a_2, \dots, a_n\}$  and  $\{b_1, b_2, \dots, b_m\}$  be generators for the ideals  $A$  and  $B$  respectively, then the product ideal  $AB$  is generated as ideal by the  $nm$  products  $a_i b_j$ , that is, every element of  $AB$  is a linear combination of these products with coefficients in the ring.

Therefore we have

$$\begin{aligned}
AB &= \frac{\alpha\beta}{4}(2, 1 + \delta)^2 \\
&= \frac{\alpha\beta}{4}(4, 2 + 2\delta, -4 + 2\delta) \\
&= \frac{\alpha\beta}{4}(2) \\
&= \left(\frac{\alpha\beta}{2}\right).
\end{aligned}$$

This proves that  $AB$  is a principal ideal generated by  $\left(\frac{\alpha\beta}{2}\right)$ .

This completes the proof.

## Result

3 of 3

We have shown that the ideal  $AB$  is generated by the element  $\frac{\alpha\beta}{2}$ , hence proves that  $AB$  is principal.

## Section 5

1. a

(a)

Recall that the norm in  $\mathbb{Z}[\sqrt{-5}]$  is given by  $N(a + b\sqrt{-5}) = a^2 + 5b^2$ . Therefore, if there is a decomposition of 11 into non-unit elements  $a$  and  $b$ , i.e.  $11 = ab$ , then

$$N(a)N(b) = 121,$$

and since  $a$  and  $b$  are not units we must have  $N(a) = N(b) = 11$ . To investigate whether that is possible, suppose there were  $c$  and  $d$  such that

$$N(c + d\sqrt{-5}) = c^2 + 5d^2 = 11.$$

By positivity of square and monotonicity of squaring we see that  $d < 2$  and that  $c < 4$ . Chcecking all the possiblites shows that there are no such  $c$  and  $d$ .

Next, in order investigate whether  $(11)$  is a prime ideal, we use the the criterion of **Proposition 13.5.1**, by which it would follow that  $(11)$  is a prime ideal if we were to show that  $\mathbb{Z}[\sqrt{-5}]/(11)$  is an integral domain. In order to determine the structure of this ring a bit more explicitly first we note that  $\mathbb{Z}[\sqrt{-5}] \cong \mathbb{Z}[X]/(X^2 + 5)$  (where  $\cong$  denotes that two objects are isomorphic), and hence

$$\mathbb{Z}[\sqrt{-5}]/(11) \cong (\mathbb{Z}[X]/(X^2 + 5))/(11) \cong (\mathbb{Z}[X]/(11))/(X^2 + 5),$$

where  $\mathbb{Z}[X]/(11)$  is the ring of integers modulo 11. Now, if we prove that  $X^2 + 5$  is irreducible over that ring, that proves that  $(\mathbb{Z}[X]/(11))/(X^2 + 5)$  is a field, and therefore an integral domain too. In order to see that  $X^2 + 5$  is irreducible, it is sufficient to check that there is no element whose square is  $-5 \equiv 6 \pmod{11}$  in  $\mathbb{Z}[X]/(11)$ , and this is straightforward as there are only 10 nontrivial cases to check. ( $2^2 \equiv 4 \pmod{11}$ ,  $3^2 \equiv 9 \pmod{11}$ ,  $4^2 \equiv 5 \pmod{11}$ , ...)

(b)

First, note that

$$(14) = (2)(7),$$

so it is sufficient to factorize  $(2)$  and  $(7)$  into prime ideals. Note that in the section **Ideal Multiplication** it is shown that in  $\mathbb{Z}[\sqrt{-5}]$  we have

$$(2) = (2, 1 + \sqrt{-5})(2, 1 - \sqrt{-5}).$$

Now this computation is over if we were to prove that  $(2, 1 + \sqrt{-5})$  and  $(2, 1 - \sqrt{-5})$  are prime; first note that they're actually equal. In order to show this, let us just show that  $1 + \sqrt{-5} \in (2, 1 - \sqrt{-5})$ , as it would be completely analogous to show that  $1 - \sqrt{-5} \in (2, 1 + \sqrt{-5})$ . The inclusion we want to show is apparent from the equality

$$1 \cdot 2 + (-1) \cdot (1 - \sqrt{-5}) = 1 + \sqrt{-5},$$

and so

$$(2) = (2, 1 + \sqrt{-5})^2.$$

Next, in order to show that  $(2, 1 + \sqrt{-5})$  is prime we prove something slightly stronger: that it is maximal. Suppose there was a proper ideal  $D$  such that  $(2, 1 + \sqrt{-5}) \subset D$ , and let  $a + b\sqrt{-5} \in D$  but  $a + b\sqrt{-5} \notin (2, 1 + \sqrt{-5})$ . Observe that  $(a + b\sqrt{-5}) - b(1 + \sqrt{-5}) = a - b$ . Now, if  $a - b$  is even, i.e.  $a - b = 2k$  for some integer, then we have

$$k \cdot 2 + b \cdot (1 + \sqrt{-5}) = a + b\sqrt{-5} \in (2, 1 + \sqrt{-5}).$$



Therefore, if  $(2, 1 + \sqrt{-5})$  were to be a proper subset of  $D$  we must have  $a - b$  odd,  $a - b = 2k + 1$ , but in this case

$$(a + b\sqrt{-5}) - (k \cdot 2 + b \cdot (1 + \sqrt{-5})) = 1 \in D$$

and therefore  $D = \mathbb{Z}[\sqrt{-5}]$ , showing that  $(2, 1 + \sqrt{-5})$  is indeed a maximal ideal, and therefore also a prime one.

### (b) (continued)

Now it would be straightforward, as in the **Ideal Multiplication** section, to verify that

$$(7) = (7, 3 + \sqrt{-5})(7, 3 - \sqrt{-5}),$$

since

$$(7, 3 + \sqrt{-5})(7, 3 - \sqrt{-5}) = (49, 21 - 7\sqrt{-5}, 21 - 7\sqrt{-5}, 14).$$

(3 was chosen as the smallest choice so that the 'last number' in the above product would be divisible by 7, i.e. smallest number of the form  $a^2 + 5$  divisible by 7.) Now we could mimick the above procedures (using the coprimality of 3 and 7) in order to show that  $(7, 3 + \sqrt{-5})$  and  $(7, 3 - \sqrt{-5})$  are prime ideals, hence giving the decomposition

$$(14) = (2, 1 + \sqrt{-5})^2(7, 3 + \sqrt{-5})(7, 3 - \sqrt{-5}).$$

### Result

4 c

In **(a)** we show that 11 is an irreducible element and that  $(11)$  is a prime ideal, while in **(b)** we arrive at decomposition  $(14) = (2, 1 + \sqrt{-5})^2(7, 3 + \sqrt{-5})(7, 3 - \sqrt{-5})$ .

## 2. a

### (a)

Recall that

$$\mathbb{Z}[\sqrt{-3}] \cong \mathbb{Z}[x]/(x^2 + 3).$$

Writing this isomorphism out we would obtain that it maps 2 to 2 and that it maps  $1 + \sqrt{-3}$  to  $x + 1$  (this is a consequence of  $\sqrt{-3}$  being mapped to  $x$ ).

Therefore

$$\begin{aligned} \mathbb{Z}[\sqrt{-3}]/(2, 1 + \sqrt{-3}) &\cong (\mathbb{Z}[x]/(x^2 + 3))/(2, 1 + x) \\ &\cong (\mathbb{Z}[x]/(2, 1 + x))/(x^2 + 3) \\ &\cong (\mathbb{Z}[x]/(2))/(1 + x)/(x^2 + 3) \\ &\cong \mathbb{Z}/(2), \end{aligned}$$

where we could separate  $\mathbb{Z}[x]/(2, 1 + x)$  as  $(\mathbb{Z}[x]/(2))/(1 + x)$  because  $\gcd(2, 1 + x) = 1$  in  $\mathbb{Z}[x]$ . Therefore, as this quotient is a field, this proves that  $(2, 1 + \sqrt{-3})$  is maximal.



(b)

We have

$$\begin{aligned}\overline{A}A &= \overline{(2, 1 + \sqrt{-3})}(2, 1 + \sqrt{-3}) = (2, 1 - \sqrt{-3})(2, 1 + \sqrt{-3}) \\ &= (4, 2 + 2\sqrt{-3}, 2 - 2\sqrt{-3}, 4) \\ &= (4, 2 + 2\sqrt{-3}, 2 - 2\sqrt{-3}) \\ &= (4, 2 + 2\sqrt{-3}) \\ &= 2(2, 1 + \sqrt{-3}) = 2A,\end{aligned}$$

where the fourth equality follows from  $2 - 2\sqrt{-3} = 4 + (-1)(2 + 2\sqrt{-3})$ . Suppose now  $\overline{A}A$  was principal, say  $\overline{A}A = (a)$ . It is easy to see from the equality  $(a) = (2)(2, 1 + \sqrt{-3})$  that if  $a$  were an integer then  $a|2$ , but then the only choice is  $a = 2$ , which would imply (as principal ideals are cancellable) that  $(2, 1 + \sqrt{-3}) = R$ , which is false. Therefore the Main Lemma does not hold. Now we could similarly as we proved that  $\overline{A}A$  is not a principal integer ideal, prove that it is not a principal ideal in general.

(c)

The inclusion  $(2) \subset A$  is straightforward as  $2 \in A$  by its definition. Now suppose that  $A$  divides  $(2)$ , i.e. that there is a proper ideal  $B$  such that  $AB = (2)$ . First note that we have  $\overline{A} = A$  since  $1 - \sqrt{-3} = 2 - (1 + \sqrt{-3})$ . Now, by (b) we have

$$2AB = \overline{A}AB = 2\overline{A} = 2A,$$

so that

$$AB = A.$$

But this implies that  $A = (2)$ , which is a contradiction.

## Result

4 of 4

In (a) we show that  $A$  is principal by showing that  $R/A$  is a field, while in (b) we obtain the equality  $\overline{A}A = 2A$  and use this to show that the Main Lemma does not hold and that  $\overline{A}A$  is not a principal ideal. In (c) we use the (b) part to derive a contradiction. Click to see more details.

## 3. a

It is sufficient to show that  $f = y^2 - x^3 - x$  is an irreducible element over  $\mathbb{C}[x, y]$  -- from there it follows that  $(f)$  is a prime ideal, and hence  $\mathbb{C}[x, y]/(f)$  is an integral domain.

In order to show that  $f$  is irreducible, we view the ring  $\mathbb{C}[x, y]$  as a polynomial ring  $R[y]$  in one variable, where  $R = \mathbb{C}[x]$ . Therefore, the polynomial in question is

$$f(y) = y^2 - x^3 - x,$$

and we can apply the Eisenstein's criterion to it. Recall that Eisenstein's criterion says that if for a polynomial  $f$  there is a prime  $p$  such that  $p$  divides every coefficient but the leading coefficient (which is 1 in  $f(y)$ ), and  $p^2$  does not divide the free coefficient (which is  $-x^3 - x$  in  $f(y)$ ), then the polynomial  $f$  is irreducible.

We propose that such a prime is given by  $p = x - i$ . First, in order to check that it is prime, we note that  $\mathbb{C}[x]/(x - i) \cong \mathbb{C}$ , which is an integral domain, and hence  $x - i$  is prime. It obviously does not divide the leading coefficient (1), and it divides the free coefficient, for  $-x^3 - x = -x(x + i)(x - i)$  over  $\mathbb{C}[x]$ .

It remains to check that  $(x + i)^2$  does not divide  $-x^3 - x$ . This follows from  $\mathbb{C}[x]$  being a unique factorization domain, but it is simple enough showing directly by supposing there was a polynomial  $f(x) = ax + b$  (as it has to be of degree 1) such that

$$(ax + b)(x + i)^2 = -x^3 - x$$

which means that

$$ax^3 + (2ia + b)x^2 - (a + 2ib)x - b = -x^3 - x,$$

where comparison of leading and free coefficient force  $b = 0$  and  $a = -1$ , but then the coefficients of  $x^2$  and  $x$  on the left-hand side and the right-hand side do not agree.

As we have verified Eisenstein's criterion, this means that the polynomial is indeed irreducible, and therefore the principal ideal is prime and therefore the ring  $\mathbb{C}[x, y]/(f)$  is an integral domain.

## Result

3 of 3

We show that it is an integral domain by showing that  $f$  is an irreducible element over  $\mathbb{C}[x, y]$ ; we do this by reducing it to the question of showing the irreducibility of the polynomial  $f(y) = y^2 - x^3 - x$  for polynomial ring  $\mathbb{C}[x][y]$ , whereby Eisenstein's criterion is applicable. Click to see more details.

## Section 6

### 1. a

Consider the provided statement to decide  $p$  splits or ramifies in  $R$  and also find a lattice basis for a prime ideal factor of  $p$ .

[Comment](#)

### Step 2 of 4 ^

As it is assumed that  $d = -14$  and  $p = 2, 3, 5, 7, 11, 13$ . By proposition 11.9.3,  $p$  stays prime in  $R$  if and only if  $x^2 + 14$  is irreducible in  $\text{mod } p$ .

If  $p = 2$  then,

$$\begin{aligned} x^2 &\equiv 0 \\ &\equiv x \cdot x \pmod{2} \end{aligned}$$

In fact (2) ramifies as,

$$(2) = (2, \sqrt{-14})(2, \sqrt{-14})$$

If  $p = 3$  then,

$$\begin{aligned} x^2 + 2 &\equiv 0 \\ &\equiv (x+1) \cdot (x+2) \pmod{3} \end{aligned}$$

3 can be ramifies as below;

$$(3) = (3, 1 + \sqrt{-14})(3, 1 - \sqrt{-14})$$

If  $p = 5$  then,

$$\begin{aligned} x^2 + 3 &\equiv 0 \\ &\equiv (x+1) \cdot (x+4) \pmod{5} \end{aligned}$$

In fact (5) ramifies as,

$$5 = (5, 1 + \sqrt{-14})(5, 1 - \sqrt{-14})$$

If  $p = 7$  then,

$$\begin{aligned} x^2 + 2 &\equiv 0 \\ &\equiv x \cdot x \pmod{7} \end{aligned}$$

In fact (7) ramifies as below;

$$(7) = (7, \sqrt{-14})(7, -\sqrt{-14})$$

If  $p = 11$  then,  $x^2 + 3 \equiv 0$

The obtained result is irreducible by inspecting  $1^2$  through  $6^2$  therefore (11) stays prime.

If  $p = 13$  then,

$$\begin{aligned} x^2 + 1 &\equiv 0 \\ &\equiv (x+5) \cdot (x+8) \pmod{13} \end{aligned}$$

In fact (13) ramifies as below;

$$(13) = (13, 5 + \sqrt{-14})(13, 5 - \sqrt{-14})$$

These ideal equivalences on the right and this can be done by the method problem 11.8.6 and it is also noticed that there is a relation between the factors of  $x^2 + 14$  and the lattice basis of the ideal factorization.

## 2. a

By the **Theorem 13.6.1 (d)** we see that 2 remains a prime if and only if the polynomial  $g(x) = x^2 - x + \frac{1-d}{4}$  is irreducible in  $\mathbb{F}_2[x]$ . Note first that we have

$$x^2 - x + a = x^2 - x + 1$$

in  $\mathbb{F}_2[x]$  if  $a$  is odd, where we observe that  $x^2 - x + 1$  is irreducible in  $\mathbb{F}_2[x]$ , and

$$x^2 - x + a = x^2 - x = x(x-1)$$

in  $\mathbb{F}_2[x]$  if  $a$  is even, so that  $g(x)$  is irreducible if and only if  $\frac{1-d}{4}$  is an odd number.

If  $\frac{1-d}{4} = 2k + 1$  for some integer  $k$  then

$$1 - d = 8k + 4,$$

or

$$d = -8k - 3,$$

where taking congruences modulo 8 we obtain

$$d \equiv -3 \equiv 5 \pmod{8}.$$

On the other hand, if  $\frac{1-d}{4} = 2k$  for some integer  $k$  then by repeating this procedure we obtain that

$$d \equiv 1 \pmod{8}.$$

(If  $d \not\equiv 1, 5 \pmod{8}$  then it  $d$  cannot be equal 1 modulo 4.)

## Result

3 of 3

2 remains a prime if  $d \equiv 5 \pmod{8}$  and does not remain a prime if  $d \equiv 1 \pmod{8}$ . Click to see more details.

3. a

(a)

Recall that we can write  $R$  as a quotient  $\mathbb{Z}[x]/(f(x))$  where  $f(x) = x^2 - d$  is  $d \equiv 2, 3 \pmod{4}$  and  $f(x) = x^2 - x + \frac{1}{4}(1 - d)$  if  $d \equiv 1 \pmod{4}$ . Therefore, we have a sequence of isomorphisms

$$R/(p) \cong (\mathbb{Z}[x]/(f(x)))/(p) \cong (\mathbb{Z}[x]/(p))/(f(x)).$$

We argue that  $(\mathbb{Z}[x]/(p))/(f(x))$  is a field with  $p^2$  elements. Since  $p$  does not split, then by **Theorem 13.6.1** it follows that  $f(x)$  is an irreducible element over  $\mathbb{Z}[x]/(p)$ . By **Proposition 11.8.4** it follows that the ideal  $(f(x))$  is then maximal over  $\mathbb{Z}[x]/(p)$ , so that the quotient is indeed a field. Lastly, since  $f(x)$  is of degree 2, it is a field of  $p^2$  elements.

(b)

If  $p$  splits but does not ramify, then we can write  $p = P\bar{P}$  where  $P$  is a prime ideal. Recall now that in an imaginary quadratic ring we have that if an ideal is prime, then it is maximal, so that  $P + \bar{P} = R$ . Note that since this holds, we have  $P\bar{P} = P \cap \bar{P}$ , so that we can apply Chinese remainder theorem for rings which gives us

$$R/(P\bar{P}) \cong R/P \times R/\bar{P}. \quad (1)$$

The conjugation maps also gives us the isomorphism  $R/P \cong R/\bar{P}$  so in order to finish our proof we just have to demonstrate that  $R/P \cong \mathbb{F}_p$ . Since  $P$  is prime then by the same argument as in (a) it is maximal and therefore  $R/P$  is a field. The only thing remaining is to show that  $R/P$  has  $p$  elements, but note that by (1) it follows that  $R/(P\bar{P})$  has  $|R/P|^2$  elements. However, we could mimick the argument of (a) to show that  $R/(P\bar{P}) = R/(p)$  has  $p^2$  elements, which shows that  $|R/P| = p$ , which is what we wanted to show.

## Result

3 of 3

The (a) part follows from writing out quotients out as well as using **Theorem 13.6.1**, while (b) follows from the similar reasoning as (a) with the added help of Chinese remainder theorem for rings. Click to see more details.

4. a



(a)

As per **Theorem 13.6.1**, it is sufficient to prove that if  $x^2 - d$  is irreducible modulo  $p$  then  $x^2 - x + \frac{1}{4}(1 - d)$  is irreducible modulo  $p$ . First note that a quadratic polynomial is irreducible over a field if and only if it has no roots in that field. It is easy to see that the usual quadratic formula holds over any field of characteristic  $p \neq 2$ , so we see that the solutions of  $x^2 - x + \frac{1}{4}(1 - d)$ , if there were any, would be

$$x_{1,2} = \frac{1 \pm \sqrt{1 - (1 - d)}}{2} = \frac{1 \pm \sqrt{d}}{2}.$$

But if  $x^2 - d$  is irreducible modulo  $p$  then  $\sqrt{d}$  does not exist modulo  $p$ , and therefore  $x^2 - x + \frac{1}{4}(1 - d)$  is also irreducible.

## Step 2

2 of 3

(b)

Observe that we used the hypothesis that the characteristic of the field is  $p \neq 2$ , so that the above reasoning does not hold for  $p = 2$ . In fact,  $x^2 - d$  is always reducible modulo 2, but  $x^2 - x + \frac{1}{4}(1 - d)$  is not necessarily -- we described, in **Exercise 6.2**, what are the requirements for 2 remaining prime and splitting in case  $d \equiv 1 \pmod{4}$ .

## Result

3 of 3

In (a) we use **Theorem 13.6.1** to reduce the statement to proving that if  $x^2 - d$  is irreducible modulo  $p \neq 2$  then  $x^2 - x + \frac{1}{4}(1 - d)$  is irreducible modulo  $p$ , and we use the usual quadratic formula (which holds for fields of characteristic  $\neq 2$ ) to show this. In (b) we note that  $x^2 - d$  is always reducible modulo  $p$ , but 2 does not always split. Click for more details.

5. a

An integer prime  $p$  is said to remain prime in the field if the principal ideal  $(p) = pR$  is a prime ideal, where  $R$  denotes the ring of integers over a quadratic field of integers. If this is not the case  $(p)$  is  $\overline{p}p$  where  $p$  is a prime ideal and  $\overline{p}$  is its conjugate, and in this scenario the prime splits. In particular if  $\overline{p} = p$ , the prime  $p$  is said to ramify.

(a) Let  $d \equiv 2$  or  $3 \pmod{4}$ .

Let  $p = 2$  and  $d \equiv 3 \pmod{4}$

Since  $d \equiv 3 \pmod{4}$ ,

So on applying modulo 2 following holds,

$$\begin{aligned} d &\equiv 3 \pmod{4} \\ &= 4q + 3, q \in \mathbb{Z} \\ &= 2(2q) + 2 + 1 \end{aligned}$$

Hence,

$$\begin{aligned} d &= 2(2q + 1) + 1 \\ &= 2m + 1, m \in \mathbb{Z} \\ &\equiv 1 \pmod{2} \end{aligned}$$

Now  $d \equiv 1 \pmod{2}$  implies that 2 divides all the generators of  $(2)$ .

So choose  $(2, 1 - \sqrt{d})$  and evaluate its square.

$$\begin{aligned} (2, 1 - \sqrt{d})^2 &= (2, 1 - \sqrt{d})(2, 1 - \sqrt{d}) \\ &= (4, 2 - 2\sqrt{d}, 1 + d - 2\sqrt{d}) \end{aligned}$$

$$(2, 1 - \sqrt{d})^2 = (2)$$

Since  $(2 - 2\sqrt{d}) - (1 + d - 2\sqrt{d}) = 1 - d$ ,

And since  $d \equiv 3 \pmod{4}$  so,

$$\begin{aligned} 1 - d &= 1 - 4q - 3 \\ &= -4q - 2 \\ &\equiv 2 \pmod{4} \end{aligned}$$

Thus  $p = 2$  ramifies in  $R$ .

Let if  $p$  divides  $d$  that is  $p \mid d$ .

Clearly the case where  $p = 2$  and  $d \equiv 2 \pmod{4}$  is covered here.

Since,

$$\begin{aligned} d &\equiv 2 \pmod{4} \\ &= 4n + 2, n \in \mathbb{Z} \\ &= 2(2n + 1) \end{aligned}$$

Here  $p = 2$  divides  $2(2n + 1)$ .

Consider  $(p, \sqrt{d})^2$ ,

$$\begin{aligned} (p, \sqrt{d})^2 &= (p, \sqrt{d})(p, -\sqrt{d}) \\ &= (p^2, p\sqrt{d}, d) \\ &= (p) \end{aligned}$$

Since  $p \mid d \Rightarrow p$  divides all the generators, and  $d$  is square free.

So,

$$\gcd(d, p^2) = p,$$

which means that there is a integer linear combination such that

$$p \in (p, \sqrt{d})^2.$$

Conversely, let there exists  $p \neq 2$  such that

$$\begin{aligned} (p) &= (p, a + b\sqrt{d})^2 \\ &= (p^2, p(a + b\sqrt{d}), a^2 - 2ab\sqrt{d} + b^2d) \\ &= A \end{aligned}$$

Then  $A \subset (p)$  only if  $p$  divides all the generators, so  $p \mid a^2 + b^2d$  and  $p \mid 2ab$ .

Since  $p \neq 2$  so  $p \nmid b$  or  $p \nmid a$ .

If  $p \nmid a$ ,  $p \mid a^2 + b^2d$  then  $p \mid b^2d$  which implies  $p \mid b$  or  $p \mid d$ . And  $p \nmid b$  implies  $p \mid a$ .

But  $p \nmid a$  and  $p \nmid b$  implies that  $p^2$  divides all the generators and thus  $p \notin A$ .

This is a contradiction.

Thus  $p \mid a$  and  $p \mid d$  must be the case.

**Therefore, the given result is proved.**

6. a



If  $p = a^2 - b^2d$  for integer  $a, b$ , then  $p = (a - b\sqrt{d})(a + b\sqrt{d})$  is a factorization of  $p$  in the ring of integers of  $\mathbb{Q}[\sqrt{d}]$ , so that

$$(p) = (a - b\sqrt{d})(a + b\sqrt{d}).$$

Therefore,  $p$  does not remain a prime, i.e. it splits. Since by **Theorem 13.6.1** every principal ideal  $(p)$  which is not prime is a product of a prime ideal and its conjugate, and by uniqueness of factorization into prime factors, it follows that  $P = (a + b\sqrt{d})$  and  $\bar{P} = (a - b\sqrt{d})$  (for if one of them was not prime, we could factorize them and hence would end with factorization of  $(p)$  in more than two prime ideals).

## Result

2 of 2

We show that in that case  $p$  splits and  $(p) = P\bar{P}$  with  $P = (a + b\sqrt{d})$ ,  $\bar{P} = (a - b\sqrt{d})$ . Click to see more details.

## 7. a

To prove that  $(p, a + \delta)$  is a lattice basis for a prime ideal that divides  $(p)$ ,

Suppose  $R$  be a ring of  $\mathbb{Q}[\delta]$  be the quadratic number field and  $p$  be a prime number other than 2.

That is,

$$p \neq 2$$

And

$$a^2 \equiv d \pmod{p}, d \equiv 2 \text{ or } 3 \pmod{4}$$

First, find the norm of the ideal  $(p, a + \delta)$ .

Suppose  $\{p, a + \delta\}$  be the lattice basis of the ideal  $(p, a + \delta)$  and

$$A = (p, a + \delta)$$

Since,

$$\delta p = p(a + \delta) - ap$$

And

$$a\delta + d = a(a + \delta) - \frac{a^2 - d}{p}(p)$$

Since,

$$\bar{\delta} = -\delta$$

Then, the norm  $N(A)$  of the ideal  $(p, a + \delta)$  is,

$$\begin{aligned} N(A) &= A\bar{A} \\ &= (p, a + \delta)(p, a - \delta) \\ &= (p^2, p(a - \delta), p(p, a + \delta), a^2 - \delta^2) \end{aligned}$$

Since,

$$\delta = \sqrt{d}$$

Then,

$$\begin{aligned} N(A) &= A\bar{A} \\ &= (p, a + \delta)(p, a - \delta) \\ &= (p^2, p(a - \delta), p(p, a + \delta), a^2 - \delta^2) \\ &= (p^2, p(a - \delta), p(p, a + \delta), a^2 - d) \end{aligned}$$

Then, the number  $p^2 / N(A)$  is a rational algebraic number.

Since, rational algebraic numbers are rational numbers.

Then, the number  $p^2 / N(A)$  is a rational number.

Then,

$$p \mid N(A)$$

This implies that,

$$N(A) \in \{p, p^2\}$$

Now, take  $\alpha$  as,

$$\begin{aligned} \alpha &= \frac{p(a + \delta) - p(a - \delta)}{N(A)} \\ &= \frac{2p\delta}{N(A)} \end{aligned}$$

Then,

$$\begin{aligned} \bar{\alpha} &= \frac{p(a + \bar{\delta}) - p(a - \bar{\delta})}{N(A)} \\ &= \frac{2p\bar{\delta}}{N(A)} \end{aligned}$$

Then,  $\delta$  is algebraic integer

Now,

$$\begin{aligned} N(\alpha) &= \alpha\bar{\alpha} \\ &= \frac{2p\delta}{N(A)} \cdot \frac{2p\bar{\delta}}{N(A)} \\ &= \frac{4p^2\delta\bar{\delta}}{N(A)^2} \\ &= \frac{4p^2}{N(A)^2} N(\delta) \end{aligned}$$

Since,

$$\begin{aligned} N(\delta) &= \delta \bar{\delta} \\ &= |d| \end{aligned}$$

Then,

$$\begin{aligned} N(\alpha) &= \frac{4p^2}{N(A)^2} N(\delta) \\ &= \frac{4p^2}{N(A)^2} |d| \\ &= \frac{4p^2}{[N(A)^2 / p^2]} |d| \end{aligned}$$

Since, the number  $p^2 / N(A)$  and  $|d|$  are rational numbers.

Then, the number  $N(\alpha)$  is a rational number.

Since,

$$a^2 \equiv d \pmod{p}, d \equiv 2 \text{ or } 3 \pmod{4}$$

And

$$N(A) \in \{p, p^2\}$$

Then,  $d$  is square free and  $p$  is odd.

And

$$N(A) \neq p^2$$

Then,

$$N(A) = p$$

Then,

$$(N(A)) = A\bar{A} = (p)$$

This implies that,  $A$  is a prime ideal and it divides  $(p)$  with the lattice basis  $\{p, a + \delta\}$ .

Hence, it is proved that  $\{p, a + \delta\}$  is a lattice basis for a prime ideal that divides  $(p)$ .

## Section 7

1. a

By **Lemma 13.7.6** we have that  $N(B^2) = N(B)^2$ , where  $N(B)$  can be computed by the remark preceeding that lemma (13.7.5), i.e. we first compute the integer such that  $\bar{B}B = (n)$  :

$$\begin{aligned} \bar{B}B &= (3, 1 + \sqrt{-5})(3, 1 - \sqrt{-5}) \\ &= (9, 3 + 3\sqrt{-5}, 3 - 3\sqrt{-5}, 6) \\ &= (3)(3, 1 + \sqrt{-5}, 1 - \sqrt{-5}, 2) \\ &= (3)R = (3), \end{aligned}$$

where  $(3, 1 + \sqrt{-5}, 1 - \sqrt{-5}, 2)$  because it contains coprime elements 3 and 2, i.e. it contains  $3 - 2 = 1$  and therefore the whole  $R$ .

Now, since  $N(B^2) = 9$ , then again by **Lemma 13.7.6** we know that if  $B^2 = (\alpha)$  then  $N(\alpha) = 9$ , and thus the only candidates for  $\alpha = a + b\sqrt{-5}$  are solutions to  $a^2 + 5b^2 = 9$ . We note that there are 4 solutions if we discard those which are integer multiples of another solution:

$$(a, b) \in \{(3, 0), (0, 3), (2, 1), (2, -1)\},$$

corresponding to numbers

$$3, 3\sqrt{-5}, 2 + \sqrt{-5}, 2 - \sqrt{-5},$$

where we can note that  $(3) = (3\sqrt{-5})$ . Also note that  $2 + \sqrt{-5} \notin (2 - \sqrt{-5})$ . Suppose to the contrary, then there's a  $\beta \in R$  such that  $2 + \sqrt{-5} = \beta(2 - \sqrt{-5})$ . We can solve that for  $\beta$  to obtain

$$\beta = \frac{2 + \sqrt{-5}}{2 - \sqrt{-5}} = \frac{(2 + \sqrt{-5})^2}{(2 - \sqrt{-5})(2 + \sqrt{-5})} = \frac{-1}{9} + \frac{4}{9}\sqrt{-5},$$

so that  $\beta \notin R$ . Similarly, as we have

$$\frac{2 - \sqrt{-5}}{2 + \sqrt{-5}} = \frac{-1}{9} + \frac{-4}{9}\sqrt{-5},$$

then  $2 - \sqrt{-5} \notin (2 + \sqrt{-5})$ .

Let us now compute  $B^2$ ,

$$B^2 = (3, 1 + \sqrt{-5})^2 = (9, 3 + 3\sqrt{-5}, -4 + 2\sqrt{-5}),$$

which allows us to eliminate the choice of 3, so that the only question left is whether  $B^2 = (2 + \sqrt{-5})$  or  $B^2 = (2 - \sqrt{-5})$ . By the result of the previous paragraph, it is sufficient to show that one of them is in  $B^2$ . Now we just notice that

$$9 + (-4 + 2\sqrt{-5}) + (-1)(3 + 3\sqrt{-5}) = 2 - \sqrt{-5},$$

to conclude that

$$B^2 = (2 - \sqrt{-5}).$$

## Result

4 of 4

We use the results on the norm of a (principal) ideal in order to show that  $B^2 = (2 - \sqrt{-5})$ . Click for more details.

## 2. a

Suppose that  $A$  and  $A'$  are similar, we want to prove there is a nonzero ideal  $C$  such that  $AC$  and  $A'C$  are both principal ideals.

$A$  and  $A'$  being similar means there is a complex number  $\lambda$  such that

$$A = \lambda A'. \tag{1}$$

Now if we multiply both sides of (1) by  $\overline{A}$  and use that  $\overline{A}A = (n)$ , we get

$$(n) = A'(\lambda \overline{A}).$$

If we multiply  $A$  by  $(\lambda \overline{A})$  we obtain

$$A(\lambda \overline{A}) = (\lambda n),$$

so that we see that the ideal  $C = \lambda \overline{A}$  fits our requirements.

Conversely, suppose that there is an ideal  $C$  such that  $AC$  and  $A'C$  were both principal, i.e.  $AC = (\lambda)$  and  $A'C = (\delta)$  for some  $\lambda$  and  $\delta$  in the ring of integers of that imaginary quadratic field. If we multiply both sides of  $AC = (\lambda)$  by  $A'$  we obtain

$$A(A'C) = \lambda A',$$

where we can substitute  $A'C = (\delta)$  in order to obtain

$$\delta A = \lambda A'. \quad (2)$$

Recall now that the definition of similarity only calls for the number with which we multiply the other ideal to be *complex*, i.e. it does not have to be in the ring of integers. Therefore, we can divide (2) through with  $\delta$  in order to obtain

$$A = \left(\frac{\lambda}{\delta}\right) A',$$

i.e.  $A$  and  $A'$  are similar.

## Result

3 of 3

Both directions follow from the definition of similarity and using the result that  $A\bar{A}$  is a principal ideal for any ideal  $A$  in the ring of integers of an imaginary quadratic number field. Click to see more details.

### 3. a

Note that  $-26 \equiv 2 \pmod{4}$ , and therefore the ring of integers in question is  $\mathbb{Z}[\sqrt{-26}]$ , and the norm of an element  $\alpha = a + b\sqrt{-26}$  is  $N(\alpha) = a^2 + 26b^2$ , so that an integer  $n$  is a norm of some  $\alpha$  if and only if the equation  $a^2 + 26b^2 = n$  has solutions with  $a, b$  integers.

Observe that this is equivalent with  $n - 26b^2$  being a square of an integer for any  $b$ , which is in our case easier to check (except for the last case).

#### **n = 75**

We have that  $75 - 26b^2$  assumes the following positive values for integer  $b$ :

$$75, 49,$$

where we see that we have a square, i.e. 49. It is achieved for  $a = \pm 7$  and  $b = \pm 1$ , so that the elements of  $R$  with norm equal to 75 are

$$\pm 7 \pm \sqrt{-26}.$$

#### **n = 250**

We find that  $250 - 26b^2$  assumes the following positive values for integer  $b$ :

$$250, 224, 146, 16,$$

where we see that 16 is a square, achieved by  $a = 4$  and  $b = 3$ . Therefore the elements of  $R$  with norm equal to 250 are

$$\pm 4 \pm 3\sqrt{-26}.$$

$$n = 375$$

We see compute that  $375 - 26b^2$  assumes the following positive values for integer  $b$ :

$$375, 349, 271, 141,$$

where we find that none of them are squares. Therefore, there is no  $\alpha \in R$  such that  $N(\alpha) = 375$ .

$$n = 5^6$$

As  $5^6$  is a square in  $\mathbb{Z}$ , we easily see that  $\alpha = 5^3 + 0\sqrt{-26} = 5^3$  fits the bill.

## Result

2 of 2

For  $n = 75$ , we find elements  $\pm 7 \pm \sqrt{-26}$ , for  $n = 250$  we find elements  $\pm 4 \pm 3\sqrt{-26}$ , for  $n = 375$  we find there are no such elements, and for  $n = 5^6$  we find  $\alpha = 5^3$ . Click to see more details.

## 4. a

(a)

To prove that the lattices  $P = (2, \delta)$  and  $Q = (3, \delta)$  are prime ideals,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $\delta^2 = -6$ .

Since,

$$\begin{aligned} 2 \cdot 3 &= 6 \\ &= -\sqrt{-6} \cdot \sqrt{-6} \end{aligned}$$

Then,  $(2, \sqrt{-6})$  and  $(3, \sqrt{-6})$  are ideals.

Then, the polynomials corresponding to the ideals are irreducible.

Then,  $\frac{R}{(2, \sqrt{-6})}$  and  $\frac{R}{(3, \sqrt{-6})}$  are fields.

Since, every field is an integral domain.

Then, the quotient ring  $\frac{R}{(2, \sqrt{-6})}$  and  $\frac{R}{(3, \sqrt{-6})}$  are integral domain.

Since, any quotient ring  $\frac{R}{I}$  is an integral domain if and only if ideal  $I$  is prime ideal.

Therefore, the ideals  $(2, \sqrt{-6})$  and  $(3, \sqrt{-6})$  are prime ideals.

Since,  $P = (2, \delta)$ ,  $Q = (3, \delta)$  and  $\delta^2 = -6$ .

Then, the lattices  $P$  and  $Q$  are prime ideals.

Hence, it is proved that the lattices  $P = (2, \delta)$  and  $Q = (3, \delta)$  are prime ideals.



(b)

To factor the principal ideal  $(6)$  into the prime ideals explicitly in the ring of integers  $R$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $\delta^2 = -6$ .

$$\begin{aligned}(6) &= (2)(3) \\ &= (\sqrt{-6})(-\sqrt{-6})\end{aligned}$$

Then,  $(2, \sqrt{-6})$  and  $(3, \sqrt{-6})$  are ideals.

Then, the polynomials corresponding to the ideals are irreducible.

Then,  $\frac{R}{(2, \sqrt{-6})}$  and  $\frac{R}{(3, \sqrt{-6})}$  are fields.

Since, every field is an integral domain.

Then, the quotient ring  $\frac{R}{(2, \sqrt{-6})}$  and  $\frac{R}{(3, \sqrt{-6})}$  are integral domain.

Since, any quotient ring  $\frac{R}{I}$  is an integral domain if and only if ideal  $I$  is prime ideal.

Therefore, the ideals  $(2, \sqrt{-6})$  and  $(3, \sqrt{-6})$  are prime ideals.

In other words, take

$$A = (2, \sqrt{-6}), B = (3, \sqrt{-6})$$

Then,

$$A\bar{A} = 2, B\bar{B} = 3$$

Since, the norms of ideals  $A$  and  $B$  are prime.

Then, the ideals  $A$  and  $B$  are prime ideals.

Hence, the required factors of the principal ideal  $(6)$  into the prime ideals are

$$\boxed{(2, \sqrt{-6}), (3, \sqrt{-6})}.$$

(c)

To determine the class group of the ring of integers  $R$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $\delta^2 = -6$ .

Since,

$$-6 \equiv 2 \pmod{4}$$

Then,

$$\begin{aligned}\mu &= 2\sqrt{\frac{|d|}{3}} \\ &= 2\sqrt{\frac{6}{3}} \\ &= 2\sqrt{2}\end{aligned}$$

Then,

$$\begin{aligned}\lfloor \mu \rfloor &= \lfloor 2\sqrt{2} \rfloor \\ &= \lfloor 2.828 \rfloor \\ &= 2\end{aligned}$$

Then,

$$d = -6, \lfloor \mu \rfloor = 2$$

Then, the class group of  $R$  is  $C_2$ .

Hence, the required class group of the ring of integers  $R$  is  $\boxed{C_2}$ .

## Section 8

1. a

As we are working in the ring of integers of  $\mathbb{Q}[\sqrt{-26}]$ , since  $-26 \equiv 2 \pmod{4}$ , it is  $\mathbb{Z}[\sqrt{-26}]$  and the associated norm is  $N(a + b\sqrt{-26}) = a^2 + 26b^2$  with  $a, b$  integers.

Therefore, it's easy to find an element with norm  $3^2 \cdot 5^2$ , as it is just 15. In order to find an element with norm  $2 \cdot 5^3 = 250$ , we have to solve the equation

$$a^2 + 26b^2 = 250.$$

But note that this was already investigate in exercise 7.3, where we found that

$$\pm 4 \pm 3\sqrt{-26}$$

have norm 250.

### Result

2 of 2

We have  $N(15) = 3^2 \cdot 5^2$  and  $N(\pm 4 \pm 3\sqrt{-26}) = 2 \cdot 5^3$ . Click for more details.

2. a

To explain the reason for the norms  $N(4 + \delta)$  and  $N(14 + \delta)$  that these norms don't need contradictory conclusions,

Suppose  $R$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$  and  $P, Q, R$  be the prime ideals of the ring  $R$ .

Where,  $d = -74$

Take an element  $\alpha$  in the ring of integers  $R$  as,

$$\alpha = a + b\sqrt{d}$$

Then,

$$\bar{\alpha} = a - b\sqrt{d}$$

Then,

$$\begin{aligned} \alpha \bar{\alpha} &= (a + b\sqrt{d})(a - b\sqrt{d}) \\ &= a^2 - (b\sqrt{d})^2 \\ &= a^2 - b^2(d) \\ &= a^2 + 74b^2 \end{aligned}$$

Then, the norm of the element  $\alpha$  of the ring  $R$  is,

$$\begin{aligned} n &= N(\alpha) \\ &= \alpha \bar{\alpha} \end{aligned}$$

Then,

$$\begin{aligned} n &= N(\alpha) \\ &= \alpha \bar{\alpha} \\ &= a^2 + 74b^2 \end{aligned}$$

Then,

$$\begin{aligned} N(4 + \delta) &= 4^2 + 74(1) \\ &= 90 \\ &= 2 \cdot 3^2 \cdot 5 \end{aligned}$$

Then, the prime ideals  $P, Q, R$  of the ring  $R$  are such that,

$$\langle P \rangle \langle Q \rangle^2 \langle S \rangle = 1$$

Where,  $\overline{P}P = (2), \overline{Q}Q = (3), \overline{S}S = (5)$

Since,  $\langle P \rangle$  and  $\langle S \rangle$  are principal ideals and

$$\langle P \rangle = 1, \langle Q \rangle^2 = 1, \langle S \rangle = 1$$

Then, by the lemma, there exists an element  $\alpha$  of the ring  $R$  with norm  $p^i q^j s^k$ .

Then, the norm of the element  $\alpha$  of the ring  $R$  is,

$$\begin{aligned} N(\alpha) &= 2 \cdot 3^2 \cdot 5 \\ &= 90 \end{aligned}$$

And

$$N(\alpha) = a^2 + 74b^2$$

Then, from the above two relations,

$$a^2 + 74b^2 = 90$$

Then, the integer solution of the above equation is,

$$a = \pm 4, b = \pm 1$$

Then, put the value of  $a$  and  $b$  in  $\alpha = a + b\sqrt{d}$ .

Then,

$$\begin{aligned} \alpha &= a + b\sqrt{d} \\ &= \pm 4 \pm \sqrt{-26} \end{aligned}$$

Then, the norm  $N(4 + \delta)$  does not lead to contraction because it is norm of an element of the ring of integers  $R$ .

In the similar way

$$\begin{aligned} N(14 + \delta) &= 14^2 + 74(1) \\ &= 270 \\ &= 2 \cdot 3^3 \cdot 5 \end{aligned}$$

Where,  $\overline{P}P = (2), \overline{Q}Q = (3), \overline{S}S = (5)$

Since,  $\langle P \rangle$  and  $\langle S \rangle$  are principal ideals and

$$\langle P \rangle = 1, \langle Q \rangle^3 = 1, \langle S \rangle = 1$$

Then, by the lemma, there exists an element  $\alpha$  of the ring  $R$  with norm  $p^i q^j s^k$ .

Then, the norm of the element  $\alpha$  of the ring  $R$  is,

$$\begin{aligned} N(\alpha) &= 2 \cdot 3^3 \cdot 5 \\ &= 270 \end{aligned}$$

And

$$N(\alpha) = a^2 + 74b^2$$

, Then, from the above two relations,

$$a^2 + 74b^2 = 270$$

Then, the integer solution of the above equation is,

$$a = \pm 14, b = \pm 1$$

Then, put the value of  $a$  and  $b$  in  $\alpha = a + b\sqrt{d}$ .

Then,

$$\begin{aligned}\alpha &= a + b\sqrt{d} \\ &= \pm 14 \pm \sqrt{-26}\end{aligned}$$

Then, the norm  $N(14 + \delta)$  does not lead to contraction because it is norm of an element of the ring of integers  $R$ .

Hence, it concludes that the norms  $N(4 + \delta)$  and  $N(14 + \delta)$  norms do not need contradictory conclusions in the ring of integers  $R$  in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

### 3. a

To compute the norms  $N(1 + \delta), N(4 + \delta), N(5 + \delta), N(9 + 2\delta)$  and  $N(11 + 2\delta)$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta^2 = -29$ .

Since, the norm of the element  $\alpha$  of the ring  $R$  is defined as,

$$N(\alpha) = \alpha \bar{\alpha}$$

Then,

$$\begin{aligned}N(1 + \delta) &= (1 + \delta)(1 - \delta) \\ &= (1 - \delta^2) \\ &= 1 + 29 \\ &= 30\end{aligned}$$

And

$$\begin{aligned}N(4 + \delta) &= (4 + \delta)(4 - \delta) \\ &= (16 - \delta^2) \\ &= 16 + 29 \\ &= 45\end{aligned}$$

And

$$\begin{aligned}N(5 + \delta) &= (5 + \delta)(5 - \delta) \\ &= (25 - \delta^2) \\ &= 25 + 29 \\ &= 54\end{aligned}$$

And

$$\begin{aligned}N(9 + 2\delta) &= (9 + 2\delta)(9 - 2\delta) \\ &= (81 - 4\delta^2) \\ &= 81 + 4 \times 29 \\ &= 197\end{aligned}$$

And

$$\begin{aligned}N(11 + 2\delta) &= (11 + 2\delta)(11 - 2\delta) \\ &= (121 - 4\delta^2) \\ &= 121 + 4 \times 29 \\ &= 237\end{aligned}$$

Hence, the required norms are  $\boxed{N(1 + \delta) = 30, N(4 + \delta) = 45, N(5 + \delta) = 54, N(9 + 2\delta) = 197, N(11 + 2\delta) = 237}$ .

To explain the conclusion about the ideals in the ring of integers  $R$ ,

Since,

$$\begin{aligned} N(1+\delta) &= 30 \\ &= 2 \cdot 3 \cdot 5 \end{aligned}$$

Since, the polynomial corresponding to the field is,  $x^2 + 29$  and this is reducible in the modulo 2, 3, and 5.

Then, all these primes split.

Suppose

$$\overline{P}P = (2), \overline{Q}Q = (3), \overline{S}S = (5)$$

Then,

$$\begin{aligned} N(1+\delta) &= (1+\delta)\overline{(1+\delta)} \\ &= (1-\delta^2) \\ &= 30 \\ &= 2 \cdot 3 \cdot 5 \\ &= \overline{P}P\overline{Q}Q\overline{S}S \end{aligned}$$

And in the similar way,

$$\begin{aligned} N(4+\delta) &= 45 \\ &= 3^2 \cdot 5 \end{aligned}$$

And

$$\begin{aligned} N(5+\delta) &= 54 \\ &= 2 \cdot 3^2 \end{aligned}$$

And

$$N(9+2\delta) = 197$$

And

$$N(11+2\delta) = 237$$

Then, the ideals of the ring of integers corresponding to the norms are prime ideals the ring.

Hence, **it is concluded that the ideals corresponding to the norms are prime ideals in the ring  $R$ .**

To determine the class group of the ring of integers  $\mathbb{Z}[\delta]$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta^2 = -29$ .

Since,

$$-29 \equiv 1 \pmod{4}$$

Then,

$$\begin{aligned} \mu &= \sqrt{\frac{-29}{3}} \\ &= \sqrt{\frac{29}{3}} \\ &= 3.109 \end{aligned}$$

Then,

$$\begin{aligned} [\mu] &= [3.109] \\ &= 3 \end{aligned}$$

Then,

$$d = -29, [\mu] = 3$$

Hence, the required class group of the ring of integers  $R$  is  $[C_3]$ .

4. a

We need to prove that, for each of those values, that the class group is trivial. The case  $d = -163$  is already done in the text (**Example 13.8.2**) so we deal with the remaining values in a similar way.

## Step 2

2 of 7

### $d = -1$

Observe that  $-1 \equiv 3 \pmod{4}$ , therefore  $\mu = \left\lfloor 2\sqrt{\frac{1}{3}} \right\rfloor = 1$ . Now by **Theorem 13.7.10** it follows that the class group is generated by the classes of prime ideals whose norms are prime integers  $p \leq 1$  -- but as there are no such primes, it follows that the class group is trivial, which shows that the unique factorization follows.

### $d = -2, -3, -7, -11$

Observe that

- $-2 \equiv 2 \pmod{4}$ , therefore  $\mu = \left\lfloor 2\sqrt{\frac{2}{3}} \right\rfloor = 1$
- $-3 \equiv 1 \pmod{4}$ , therefore  $\mu = \left\lfloor \sqrt{\frac{3}{3}} \right\rfloor = 1$ ,
- $-7 \equiv 1 \pmod{4}$ , therefore  $\mu = \left\lfloor \sqrt{\frac{7}{3}} \right\rfloor = 1$ ,
- $-11 \equiv 1 \pmod{4}$ , therefore  $\mu = \left\lfloor \sqrt{\frac{11}{3}} \right\rfloor = 1$ ,

so the analogous reasoning as for  $d = -1$  shows that these groups are trivial, and hence the uniqueness factorization holds.

### $d = -19$

Observe that  $-19 \equiv 1 \pmod{4}$ , therefore  $\mu = \left\lfloor \sqrt{\frac{19}{3}} \right\rfloor = 2$ . We need to check whether the prime 2 splits in the ring of integers of  $\mathbb{Q}[\sqrt{-19}]$ . Recall that by **Theorem 13.6.1**, 2 remains a prime if and only if the polynomial  $x^2 - x + \frac{1}{4}(1 - (-19)) = x^2 - x + 5$  is irreducible modulo 2.

It is easy to check that it is indeed irreducible, for  $x^2 - x + 5 \equiv x^2 - x + 1 \pmod{2}$ , and since quadratic polynomials are irreducible if and only if they don't have any roots, it suffices to check that  $0^2 - 0 + 1 \equiv 1 \pmod{2}$  and that  $1^2 - 1 + 1 \equiv 1 \pmod{2}$ .

## Step 5

5 of 7

### $d = -43$

Observe that  $-43 \equiv 1 \pmod{4}$ , therefore  $\mu = \left\lfloor \sqrt{\frac{43}{3}} \right\rfloor = 3$ , so we have to check whether 2 and 3 remain prime in the ring of integers of  $\mathbb{Q}[\sqrt{-43}]$ . Similarly as in the  $d = -19$  case, we have to check whether  $x^2 - x + \frac{1}{4}(1 - (-43)) = x^2 - x + 11$  is irreducible modulo 2 and 3.

The case modulo 2 is analogous as before, so let us note that  $x^2 - x + 11 \equiv x^2 + 2x + 2 \pmod{3}$ , wherein we check whether that polynomial has any roots modulo 3:  $0^2 + 0 + 2 \equiv 2 \pmod{3}$ ,  $1^2 + 2 + 2 \equiv 2 \pmod{3}$ ,  $2^2 + 4 + 2 \equiv 1 \pmod{3}$  -- showing that  $x^2 - x + 11$  is indeed irreducible modulo 3.



**d = -67**

Observe that  $-67 \equiv 1 \pmod{4}$ , therefore  $\mu = \left\lfloor \sqrt{\frac{67}{3}} \right\rfloor = 4$ , and therefore we have to check whether  $x^2 - x + \frac{1}{4}(1 - (-67)) = x^2 - x + 17$  remains irreducible modulo 2 and 3. Modulo 2 is the same story as before, while for modulo 3 we note that  $x^2 - x + 17 \equiv x^2 + 2x + 2 \pmod{3}$ , and so this case too is the same as for  $d = -43$ .

## Result

7 of 7

In each of these cases we show that the class group is trivial following the standard procedure outlined in the text.  
[Click to see more.](#)

## 5. a

Determine the class group of each case and draw the possible shapes of the lattice in each case.

[Comment](#)

Step 2 of 6 ^

(a)

Consider the following case:

$$d = -10$$

Then for  $d = -10$ , to find the class group and check some conditions those are given below,

$$d \equiv 2 \text{ or } 3 \pmod{4}$$

Here,  $-10 \equiv 2 \pmod{4}$  is true, then it shows that it has an ideal class group.



(d)

Consider the following case:

$$d = -13$$

Also it shows that lattice is rectangular  $-14 \equiv 2 \pmod{4}$ .

For  $d = -13$ ,  $-13 \equiv 3 \pmod{4}$  is true, which shows that it has ideal class group, also shows that lattice is rectangular.



(c)

Consider the following case:

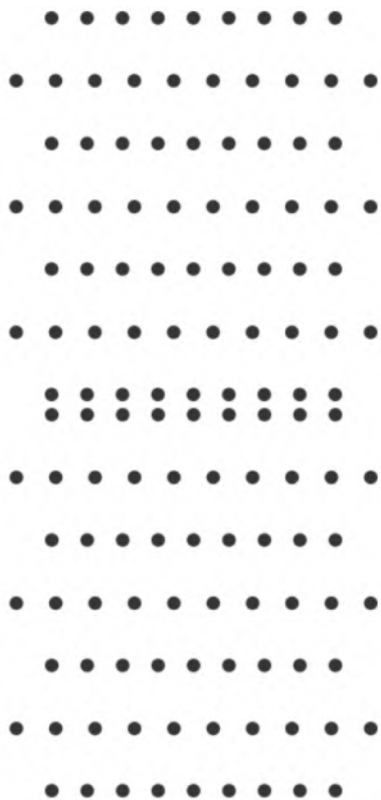
$$d = -14$$

For  $d = -14$ ,  $-14 \equiv 2 \pmod{4}$  is true,

[Comment](#)

Step 5 of 6 ^

This is shows that it has ideal class group, also shows that lattice is rectangular.



(d)

Consider the following case:

$$d = -21$$

For  $d = -21$ ,  $-21 \equiv 3 \pmod{4}$  is true, which shows that it has ideal class group, also shows that lattice is rectangular.



6. a

(a)

To determine the class group of the ring of integers  $R$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $d = -41$ .

Since,

$$\begin{aligned} -41 &= 44 - 41 \\ &\equiv 3 \pmod{4} \end{aligned}$$

Then,

$$\begin{aligned} \mu &= 2\sqrt{\frac{|d|}{3}} \\ &= 2\sqrt{\frac{41}{3}} \\ &= 7.392 \end{aligned}$$

Then,

$$\begin{aligned} \lfloor \mu \rfloor &= \lfloor 7.392 \rfloor \\ &= 7 \end{aligned}$$

Suppose  $(\alpha)$  be an ideal of the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$

Take  $\alpha = a + b\sqrt{d}$

Then,

$$\bar{\alpha} = a - b\sqrt{d}$$

Then, the product of  $\alpha$  and  $\bar{\alpha}$  is,

$$\begin{aligned}\alpha\bar{\alpha} &= (a + b\sqrt{d})(a - b\sqrt{d}) \\ &= a^2 - (b\sqrt{d})^2 \\ &= a^2 - b^2(d) \\ &= a^2 + 41b^2\end{aligned}$$

Since,

$$N(\alpha) = \alpha\bar{\alpha}$$

Then,

$$N(\alpha) = a^2 + 41b^2$$

Since,

$$\begin{aligned}N(\alpha) &\leq \lfloor \mu \rfloor \\ &= 7\end{aligned}$$

Then, to find the value of  $a$  and  $b$ .

$$\begin{aligned}N(\alpha) &= a^2 + 41b^2 \\ 7 &= a^2 + 41b^2\end{aligned}$$

Then, there is no integer solution of the above equation.

And the polynomial  $x^2 - d = x^2 + 41$  is irreducible in modulo  $p = 2, 3, 5, 7$ .

Then, the ring  $R$  is a unique factorization domain.

Since, the ring  $R$  in an imaginary quadratic number field is a principal ideal domain if and only if it is unique factorization domain.

Then, the ring  $R$  is a principal ideal domain.

Since, the class group of the ring  $R$  is defined as,

$$C(R) = \frac{\{\text{fractional ideal of } R\}}{\{\text{principal ideals of } R\}}$$

Then, class group of the ring of integers  $R$  is trivial.

Hence, the required class group of the ring of integers  $R$  in the imaginary quadratic field

$$\mathbb{Q}[\sqrt{-41}] \text{ is } [C_1].$$

(b)

To determine the class group of the ring of integers  $R$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $d = -57$ .

Since,

$$\begin{aligned}-57 &\equiv -1 \pmod{4} \\ &\equiv 3 \pmod{4}\end{aligned}$$

And

$$\begin{aligned} 57 &= 3 \cdot 19 \\ &= (-\sqrt{-57})(\sqrt{-57}) \end{aligned}$$

Then, the ideals of the ring of integers are  $(3, \sqrt{-57})$  and  $(19, \sqrt{-57})$

Then,

$$\begin{aligned} (3, \sqrt{-57})(3, \sqrt{-57}) &= (3, -\sqrt{-57})(3, \sqrt{-57}) \\ &= (9, 3 - \sqrt{-57}, 3\sqrt{-57}, 57) \end{aligned}$$

Then, it concludes that

$$(3, \sqrt{-57})(3, \sqrt{-57}) = (3)$$

In the similar way,

$$\begin{aligned} (19, \sqrt{-57})(19, \sqrt{-57}) &= (19, -\sqrt{-57})(19, \sqrt{-57}) \\ &= (19^2, 19 - \sqrt{-57}, 19\sqrt{-57}, 57) \end{aligned}$$

Then, it concludes that

$$(19, \sqrt{-57})(19, \sqrt{-57}) = (19)$$

Then, the ideals  $(3)$  and  $(19)$  are principal ideals of the ring  $R$ .

Then, the class group is  $C_2$ .

Hence, the required class group of the ring of integers  $R$  in the imaginary quadratic field

$$\mathbb{Q}[\sqrt{-41}] \text{ is } [C_2].$$

(c)

To determine the class group of the ring of integers  $R$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $d = -61$ .

Since,

$$\begin{aligned} -61 &\equiv -1 \pmod{4} \\ &\equiv 3 \pmod{4} \end{aligned}$$

Then, the polynomial  $x^2 - d = x^2 + 41$  is irreducible in modulo  $p = 2, 3, 5, 7$ .

Then, the ring  $R$  is a unique factorization domain.

Since, the ring  $R$  in an imaginary quadratic number field is a principal ideal domain if and only if it is unique factorization domain.

Then, the ring  $R$  is a principal ideal domain.

Since, the class group of the ring  $R$  is defined as,

$$C(R) = \frac{\{\text{fractional ideal of } R\}}{\{\text{principal ideals of } R\}}$$

Then, class group of the ring of integers  $R$  is trivial.

Hence, the required class group of the ring of integers  $R$  in the imaginary quadratic field

$$\mathbb{Q}[\sqrt{-61}] \text{ is trivial class group } [C_1].$$



(d)

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $d = -77$ .

Since,

$$\begin{aligned} -77 &\equiv -1 \pmod{4} \\ &\equiv 3 \pmod{4} \end{aligned}$$

And

$$\begin{aligned} 77 &= 7 \cdot 11 \\ &= (-\sqrt{-77})(\sqrt{-77}) \end{aligned}$$

Then, the ideals of the ring of integers  $R$  are  $(7, \sqrt{-77})$  and  $(11, \sqrt{-77})$

Then,

$$\begin{aligned} \overline{(7, \sqrt{-77})}(7, \sqrt{-77}) &= (7, -\sqrt{-77})(7, \sqrt{-77}) \\ &= (49, 7\sqrt{-77}, -7\sqrt{-77}, 77) \end{aligned}$$

Then, it concludes that

$$\overline{(7, \sqrt{-77})}(7, \sqrt{-77}) = (7)$$

In the similar way,

$$\begin{aligned} \overline{(11, \sqrt{-77})}(11, \sqrt{-77}) &= (11, -\sqrt{-77})(11, \sqrt{-77}) \\ &= (11^2, 11\sqrt{-77}, 11\sqrt{-77}, 77) \end{aligned}$$

Then, it concludes that

$$\overline{(11, \sqrt{-77})}(11, \sqrt{-77}) = (11)$$

Then, the ideals  $(7)$  and  $(11)$  are principal ideals of the ring  $R$ .

Then, the class group is  $C_2$ .

Hence, the required class group of the ring of integers  $R$  in the imaginary quadratic field

$$\mathbb{Q}[\sqrt{-77}] \text{ is } \boxed{C_2}.$$

(e)

To determine the class group of the ring of integers  $R$ ,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $d = -89$ .

Since,

$$\begin{aligned} -89 &\equiv -1 \pmod{4} \\ &\equiv 3 \pmod{4} \end{aligned}$$

Then, the polynomial  $x^2 - d = x^2 + 41$  is irreducible in modulo  $p = 2, 3, 5, 7$ .

Then, the ring  $R$  is a unique factorization domain.

Since, the ring  $R$  in an imaginary quadratic number field is a principal ideal domain if and only if it is unique factorization domain.

Then, the ring  $R$  is a principal ideal domain.

Since, the class group of the ring  $R$  is defined as,

$$C(R) = \frac{\{\text{fractional ideal of } R\}}{\{\text{principal ideals of } R\}}$$

Then, class group of the ring of integers  $R$  is trivial.

Hence, the required class group of the ring of integers  $R$  in the imaginary quadratic field

$$\mathbb{Q}[\sqrt{-89}] \text{ is trivial class group } \boxed{C_1}.$$

## Section 9

### 1. a

It was already shown in the text that  $1 + \sqrt{2}$  is a unit, but an even more straightforward demonstration, without using the norm, is to note that

$$(\sqrt{2} + 1)(\sqrt{2} - 1) = (\sqrt{2})^2 - 1 = 1,$$

i.e.  $1 + \sqrt{2}$  is invertible in  $\mathbb{Z}$ .

Suppose now that it is not of infinite order, so that there exists a natural number  $n$  such that

$$(1 + \sqrt{2})^n = 1.$$

But then, taking the  $n$ th root of both sides (which is fine as  $1 + \sqrt{2} > 0$ ) we obtain

$$1 + \sqrt{2} = 1,$$

which is a contradiction, and hence it has an infinite order.

### Result

2 of 2

We first show that it is a unit by showing that  $\sqrt{2} - 1$  is the inverse of  $1 + \sqrt{2}$ . Then we note that if it had finite order there would exist an  $n$  such that  $(1 + \sqrt{2})^n = 1$  and derive a contradiction from that. Click to see more details.

### 2. a

Suppose first that  $d$  is a square, i.e.  $d = k^2$  for some positive integer  $k$ . Then the equation we're asked to solve transforms into

$$(x - yk)(x + yk) = 1.$$

As both  $x - yk$  and  $x + yk$  are integers, this equation can only hold if both are equal to 1 or both are equal to  $-1$ . Suppose that  $x - yk = 1$  and  $x + yk = 1$ , then  $2x = 2$ , and thus  $x = 1$ . Using this we obtain that the only solutions are  $x = 1, y = 0$  and  $x = -1$  and  $y = 0$ . For the rest of this answer we assume that  $d$  is not a square.

### Step 2

2 of 6

Note now that  $x^2 - y^2d$  has the factorization  $(x - y\sqrt{d})(x + y\sqrt{d})$  in the ring  $\mathbb{Z}[\sqrt{d}]$ . Therefore, we are to solve

$$\alpha\bar{\alpha} = 1$$

for an  $\alpha \in \mathbb{Z}[\sqrt{d}]$ . Using the norm function  $N(a + b\sqrt{d}) = a^2 + db^2$ , we see that we are looking for all  $\alpha$  such that

$$N(\alpha) = 1.$$

As we are working in a real quadratic ring, there is a significant difference compared to working in an imaginary one, namely that we have an ordering on our elements (as they are just real numbers). Suppose then that there is the smallest  $\alpha_0 > 1$  such that  $N(\alpha_0) = 1$ . Then, for any other  $\alpha$  such that  $N(\alpha) = 1$ , there is an integer  $n$  such that

$$\alpha_0^n \leq \alpha \leq \alpha_0^{n+1}.$$

Then, defining a number  $\alpha' = \alpha \alpha_0^{-n}$ , we see clearly that  $1 \leq \alpha' < \alpha_0$ , but  $\alpha'$  is also a unit since we have

$$N(\alpha') = N(\alpha) N(\alpha_0)^{-n} = 1.$$

Now, by assumption of minimality of  $\alpha_0$ , it follows that  $\alpha' = 1$ , and therefore  $\alpha \alpha_0^{-n} = 1$ , i.e.  $\alpha = \alpha_0^n$ .

From this it follows that if we find a smallest solution  $\alpha_0 = x_0 + y_0\sqrt{d}$ , then any other solution is given by  $(x_0 + y_0\sqrt{d})^n$ ,  $n$  a positive integer. It remains to show that there is indeed a smallest solution. This is a consequence of Dirichlet's theorem, which grants us existence of certain approximations to irrational numbers. In particular, it says that if  $x$  is a real number, then for any integer  $n$  there exist integers  $p$  and  $q$  such that  $1 \leq q \leq n$  and

$$\left| x - \frac{p}{q} \right| < \frac{1}{nq}.$$

This can be proved via pigeonhole principle, by considering fractional parts  $\{0x\}, \{1x\}, \{2x\}, \dots, \{nx\}$ . As we have  $n + 1$  points in the segment  $[0, 1]$ , two of them will differ by less than  $1/n$ , say  $\{tx\}$  and  $\{sx\}$ . But then there is an integer  $p$  such that  $|(t - s)x - p| < \frac{1}{n}$ , where setting  $q = t - s$  and dividing through finishes the proof.

A direct consequence of this theorem is that there exist infinitely many integers  $(p, q)$  such that

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2}.$$

Now we are ready to prove that there is indeed a minimal solution, greater than 1, of our equation. Applying the consequence of Dirichlet's theorem to  $x = \sqrt{d}$ , we obtain infinitely many pairs  $(p, q)$  such that

$$\left| \sqrt{d} - \frac{p}{q} \right| < \frac{1}{q^2}.$$

Multiplying through by  $q$  we obtain

$$|q\sqrt{d} - p| < \frac{1}{q},$$

where we note that  $|q^2 - dp^2| = |q\sqrt{d} - p||q\sqrt{d} + p|$ .

Via triangle inequality we can bound  $|q\sqrt{d} + p| < \frac{1}{q} + 2\sqrt{d}q$ , which allows us to prove that

$$|q^2 - dp^2| < \frac{1}{q} \left( \frac{1}{q} + 2\sqrt{d}q \right) \leq 1 + 2\sqrt{d}.$$

Therefore, we have obtained infinitely many solutions  $(p, q)$  to  $|q^2 - dp^2| < 1 + 2\sqrt{d}$ , and therefore there must be an integer  $n$  such that there are infinitely many solutions  $(p, q)$  to  $q^2 - dp^2 = n$ . There are then two solutions  $(p_0, q_0)$  and  $(p_1, q_1)$ , which satisfy

$$p_0 \equiv p_1 \pmod{n} \text{ and } q_0 \equiv q_1 \pmod{n}.$$

Denoting then as  $x_0 = q_0 + p_0\sqrt{d}$  and  $x_1 = q_1 + p_1\sqrt{d}$ , without the loss of generality taking  $x_0 > x_1$ , we obtain that  $\frac{x_0}{x_1} > 1$  has norm 1 in  $\mathbb{Z}[\sqrt{d}]$ , and hence corresponds to a nontrivial solution of our equation.

## Result

6 of 6

We note that the left-hand side of our equation can be factored in  $\mathbb{Z}[\sqrt{d}]$  and use that to prove that each solution is of the form  $\alpha^n$ , where  $\alpha$  is the minimal solution in  $\mathbb{Z}[\sqrt{d}]$ . Click for more details.

### 3. a

(a)

To prove that the size function  $\sigma(\alpha) = |N(\alpha)|$  make the ring  $\mathbb{Z}[\sqrt{2}]$  into a Euclidean domain,

Suppose  $\mathbb{Z}[\sqrt{2}]$  be the ring of integers in the real quadratic field  $\mathbb{Q}[\sqrt{2}]$ .

Then, the ring  $R$ ,

$$R = \mathbb{Z}[\sqrt{2}] = \{x + y\sqrt{2} \mid x, y \in \mathbb{Z}\}$$

Then, any  $\alpha \in \mathbb{Z}[\sqrt{2}]$  such that

$$\alpha = x + y\sqrt{2}$$

And then

$$\alpha' = x - y\sqrt{2}$$

Where  $\alpha'$  is conjugate element of the element  $\alpha$ .

Then,

$$(\alpha, \alpha') \in \mathbb{R}^2$$

Then, the norm of the  $\alpha$  is defined as,

$$\begin{aligned} N(\alpha) &= \alpha' \alpha \\ &= (x + y\sqrt{2})(x - y\sqrt{2}) \\ &= x^2 - 2y^2 \end{aligned}$$

Then, the size function  $\sigma(\alpha)$  is,

$$\begin{aligned} \sigma(\alpha) &= |N(\alpha)| \\ &= |x^2 - 2y^2| \\ &= x^2 + 2y^2 \end{aligned}$$

Suppose any other  $u + \sqrt{2}v \in \mathbb{Z}[\sqrt{2}]$

Then, for some  $e$  and  $f$  such that

$$\begin{aligned} \frac{x + \sqrt{2}y}{u + \sqrt{2}v} &= \frac{(x + \sqrt{2}y)(u - \sqrt{2}v)}{(u + \sqrt{2}v)(u - \sqrt{2}v)} \\ &= \frac{(x + \sqrt{2}y)(u - \sqrt{2}v)}{(u^2 - 2v^2)} \\ &= e + \sqrt{2}f \end{aligned}$$

Now, choose  $m$  and  $n$  such that,

$$|e - m|, |f - n| \leq \frac{1}{2}$$

Suppose  $q$  be the quotient and  $r$  be the remainder,

Then, take

$$q = m + \sqrt{2}n, r = x + \sqrt{2}y - q(u + \sqrt{2}v)$$

Then,

$$x + \sqrt{2}y = q(u + \sqrt{2}v) + r$$

Then,

$$\begin{aligned} N(r) &= N\left((u + \sqrt{2}v)((e-m) + \sqrt{2}(f-n))\right) \\ &= N(u + \sqrt{2}v)N((e-m) + \sqrt{2}(f-n)) \\ &\leq \frac{3}{4}N(u + \sqrt{2}v) \\ &< N(u + \sqrt{2}v) \end{aligned}$$

This implies that,

$$\begin{aligned} \sigma(r) &= |N(r)| \\ &< |N(u + \sqrt{2}v)| \end{aligned}$$

Therefore, the ring  $\mathbb{Z}[\sqrt{2}]$  is a Euclidean domain.

Since, every Euclidean domain is a unique factorization domain.

Then, the ring  $\mathbb{Z}[\sqrt{2}]$  is also a unique factorization domain.

Hence, it is proved that the size function  $\sigma(\alpha) = |N(\alpha)|$  makes the ring  $\mathbb{Z}[\sqrt{2}]$  into a Euclidean domain and  $\mathbb{Z}[\sqrt{2}]$  is a factorization domain.

(b)

To make sketch for the principal ideal  $(\sqrt{2})$  of the ring  $\mathbb{Z}[\sqrt{2}]$ ,

Suppose  $\mathbb{Z}[\sqrt{2}]$  be the ring of integers in the real quadratic field  $\mathbb{Q}[\sqrt{2}]$ .

Since, the size function  $\sigma(\alpha)$  is defined as,

$$\sigma(\alpha) = |N(\alpha)|$$

Then, this makes the ring  $\mathbb{Z}[\sqrt{2}]$  into a Euclidean domain.

Since, every ideal in Euclidean domain is a principal ideal.

Then, the ideal  $(\sqrt{2})$  is the principal ideal in the ring  $\mathbb{Z}[\sqrt{2}]$ .

Then, any  $\alpha \in \mathbb{Z}[\sqrt{2}]$  such that

$$\alpha = x + y\sqrt{2}$$

And then

$$\alpha' = x - y\sqrt{2}$$

Then, the norm of the  $\alpha$  is defined as,

$$\begin{aligned} N(\alpha) &= \alpha'\alpha \\ &= (x + y\sqrt{2})(x - y\sqrt{2}) \\ &= x^2 - 2y^2 \end{aligned}$$

Then, for the units' in  $(u, v)$ -coordinates

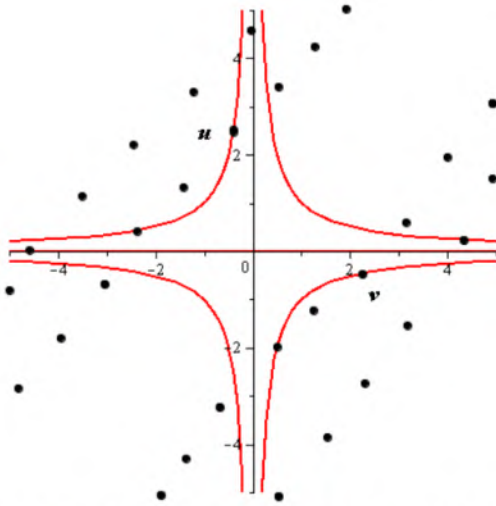
$$\begin{aligned} N(\alpha) &= \pm 1 \\ x^2 - 2y^2 &= \pm 1 \end{aligned}$$

That is, for the units of the ring  $\mathbb{Z}[\sqrt{2}]$  in  $(u, v)$ -coordinates

$$uv = \pm 1$$

Use MAPLE to sketch the principal ideal  $(\sqrt{2})$ ,





Then, the units are the points of the lattice of the principal ideal  $(\sqrt{2})$ .

Hence, above is the required sketch of the principal ideal  $(\sqrt{2})$ .

4. a

To find the possible structures for the group of units in the ring of integers  $R$ ,

Suppose  $R$  be the ring of integers in the real quadratic field  $\mathbb{Q}[\sqrt{d}]$ .

Where,  $d > 0$

Since, for  $d > 0$ , field  $\mathbb{Q}[\sqrt{d}]$  is a subfield of the field of the real numbers  $\mathbb{R}$ .

And the ring  $R$  is,

$$R = \mathbb{Z}[\sqrt{d}] = \{x + y\sqrt{d} \mid x, y \in \mathbb{Z}\}$$

Then, take

$$\alpha = x + y\sqrt{d}, \alpha' = x - y\sqrt{d}$$

Where  $\alpha'$  is conjugate element of the element  $\alpha$ .

Then,

$$(\alpha, \alpha') \in \mathbb{R}^2$$

Then, the norm of the  $\alpha'$  is defined as,

$$\begin{aligned} N(\alpha) &= \alpha'\alpha \\ &= (x + y\sqrt{d})(x - y\sqrt{d}) \\ &= x^2 - y^2d \end{aligned}$$

Since, an element  $\alpha \in R$  is a unit if and only if  $N(\alpha) = 1$ .

Then,

$$N(\alpha) = 1$$

$$\alpha'\alpha = 1$$

$$x^2 - y^2d = 1$$

This is the equation of hyperbola in the plane  $\mathbb{R}^2$ .

Since,

$$d > 0$$

Then, the equation  $x^2 - y^2d = 1$  always remains the equation of hyperbola.

Hence, the required possible structure of for the group of units in the ring of integers  $R$  is hyperbola.

5. a



(a)

To prove that the set of units  $U_0$  of the ring  $R$  is an infinite cyclic subgroup of the group of units,

Suppose  $R$  be the ring of integers in the real quadratic field  $\mathbb{Q}[\sqrt{d}]$  and  $U_0$  the set of units of the ring  $R$ .

Where,  $d > 0$

Then,

$$R = \mathbb{Z}[\sqrt{d}] = \{x + y\sqrt{d} \mid x, y \in \mathbb{Z}\}$$

And

$$U_0 = \{\alpha \in \mathbb{Z}[\sqrt{d}] \mid \alpha' \alpha = 1\}$$

Where  $\alpha'$  is conjugate element of the element  $\alpha$ .

First, show that  $U_0$  is a group under multiplication.

Suppose  $\alpha, \beta \in U_0$  such that

$$\alpha = x_1 + y_1\sqrt{d}, \beta = x_2 + y_2\sqrt{d}$$

Where,  $x_1, x_2, y_1, y_2 \in \mathbb{Z}$

Then,

$$\begin{aligned} \alpha\beta &= (x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d}) \\ &= (x_1x_2) + (y_1y_2d) + x_1y_2\sqrt{d} + x_2y_1\sqrt{d} \end{aligned}$$

Since, for  $x_1, x_2, y_1, y_2 \in \mathbb{Z}$

Then,

$$(x_1x_2) \in \mathbb{Z}, (y_1y_2d) \in \mathbb{Z}, x_1y_2 \in \mathbb{Z}, x_2y_1 \in \mathbb{Z}$$

This implies that,

$$\alpha\beta \in U_0$$

This implies that the set of units  $U_0$  of the ring is closed under multiplication.

And for  $\alpha, \beta, \gamma \in U_0$

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma$$

Then, the set of units  $U_0$  of the ring is associative under multiplication.

And for each  $\alpha \in U_0$ , there exists an element  $\alpha' \in U_0$  such that,

$$\alpha' \alpha = 1$$

Since, the set of units  $U_0$  of the ring is closed under multiplication.

Then,

$$\alpha' \alpha \in U_0$$

Then,

$$1 \in U_0$$

Or

$$1 = 1 + 0\sqrt{d} \in U_0$$

This implies that the set of units  $U_0$  of the ring has multiplicative identity.

Since, for each  $\alpha \in U_0$ , there exists an element  $\alpha' \in U_0$  such that,

$$\alpha'\alpha = 1$$

This implies that, for each element of the set of units  $U_0$  has inverse in  $U_0$ .

Hence, it concludes that, the set of units  $U_0$  forms the group under multiplication.

Then, for any  $\alpha \in U_0$

$$\alpha = x + y\sqrt{d}$$

And then

$$\begin{aligned}\alpha^2 &= (x + y\sqrt{d})^2 \\ &= (x^2 + y^2d + 2xy\sqrt{d})\end{aligned}$$

In the similar way,

$$\begin{aligned}\alpha^3 &= \alpha^2\alpha \\ &= (x^2 + y^2d + 2xy\sqrt{d})(x + y\sqrt{d}) \\ &= x^3 + xy^2d + 2x^2y\sqrt{d} + xy^2 + y^3d + 2xy^2d\end{aligned}$$

Then, there is no finite number  $n$ , for which the element  $\alpha$  has finite order.

This implies that the element  $\alpha \in U_0$  has infinite order in the group of units  $U_0$ .

And order of the group of units  $U_0$  is also infinite.

Then, the set of units  $U_0$  forms the infinite cyclic group of units.

Hence, **it is proved.**

(b)

To find the generators for the infinite cyclic group  $U_0$  of units of the ring of integers  $R$ ,

Suppose  $R$  be the ring of integers in the real quadratic field  $\mathbb{Q}[\sqrt{d}]$  and  $U_0$  infinite cyclic group  $U_0$  of units of the ring of integers  $R$ .

Where,  $d > 0$

Since, the ring  $R$  is,

$$R = \mathbb{Z}[\sqrt{d}] = \{x + y\sqrt{d} \mid x, y \in \mathbb{Z}\}$$

And the group of units  $U_0$  is,

$$U_0 = \{\alpha \in \mathbb{Z}[\sqrt{d}] \mid \alpha'\alpha = 1\}$$

Where  $\alpha'$  is conjugate element of the element  $\alpha$ .

Then, for any  $\alpha \in U_0$

$$\alpha = x + y\sqrt{d}, \alpha' = x - y\sqrt{d}$$

Then, the norm of the  $\alpha'$  is defined as,

$$\begin{aligned}N(\alpha) &= \alpha'\alpha \\ &= (x + y\sqrt{d})(x - y\sqrt{d}) \\ &= x^2 - y^2d\end{aligned}$$

Since, for any  $\alpha \in U_0$

$$\alpha'\alpha = 1$$

Then,

$$\begin{aligned}\alpha'\alpha &= x^2 - y^2d \\ 1 &= x^2 - y^2d\end{aligned}$$

That is,

$$x^2 - y^2d = 1$$

Then, for  $d = 3$ ,

$$x^2 - 3y^2 = 1$$

Then, the integer solutions of the equation are,

$$(x, y) = (2, 1), (-2, -1)$$

Put the value of  $(x, y)$  and  $d$  in  $\alpha = x + y\sqrt{d}$ .

Then,

$$\alpha = 2 + \sqrt{3}, -2 - \sqrt{3}$$

Hence, the required generators when  $d = 3$  of the infinite cyclic group  $U_0$  of units of the ring of integers  $R$  are  $\boxed{(2 + \sqrt{3}), (-2 - \sqrt{3})}$ .

Now, when  $d = 5$

Since,

$$x^2 - y^2d = 1$$

Then,

$$x^2 - 5y^2 = 1$$

Then, the integer solutions of the equation are,

$$(x, y) = (9, 4), (-9, -4)$$

Put the value of  $(x, y)$  and  $d$  in  $\alpha = x + y\sqrt{d}$ .

Then,

$$\alpha = 9 + 4\sqrt{3}, -9 - 4\sqrt{3}$$

Hence, the required generators when  $d = 5$  of the infinite cyclic group  $U_0$  of units of the ring of integers  $R$  are  $\boxed{(9 + 4\sqrt{3}), (-9 - 4\sqrt{3})}$ .

(c)

To draw the figure for the hyperbola and the units for  $d = 3$ ,

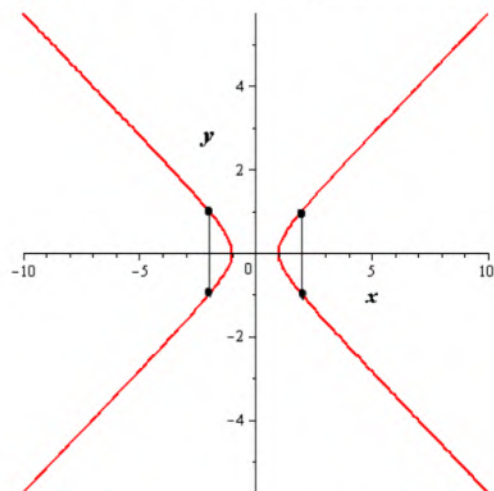
Suppose  $R$  be the ring of integers in the real quadratic field  $\mathbb{Q}[\sqrt{d}]$  and  $U_0$  infinite cyclic group  $U_0$  of units of the ring of integers  $R$ .

Since, for  $d = 3$ , the equation of hyperbola is,

$$x^2 - 3y^2 = 1$$

Use MAPLE to draw the figure for the hyperbola and the units for  $d = 3$ ,

Use MAPLE to draw the figure for the hyperbola and the units for  $d = 3$ ,



This is the required figure for the hyperbola and the units for  $d = 3$  where dots points are the units of infinite cyclic group  $U_0$ .

## Section 10

1. a

To determine the index  $[M : L]$ ,

Suppose  $M$  be the integer lattice in  $\mathbb{R}^2$  and  $L$  be the lattice with basis  $\{(2, 3)', (3, 6)'\}$ .

Since,  $M$  and  $N$  are the integer lattices in the plane  $\mathbb{R}^2$  and  $L \subset M$ , then

$$[M : L] = \frac{\Delta L}{\Delta M}$$

Then, find the  $\Delta L$  and  $\Delta M$ .

Since,  $\Delta L$  is the area of the parallelogram of the linear combination of the lattice basis vectors

$$\Pi((2, 3)', (3, 6)').$$

Then,

$$\begin{aligned}\Delta L &= \Pi((2, 3)', (3, 6)') \\ &= 3\end{aligned}$$

Then,

$$\begin{aligned}[M : L] &= \frac{\Delta L}{\Delta M} \\ &= \frac{3}{\Delta M}\end{aligned}$$

Since, the indexes  $[M : L]$  are always integers.

Then, the possibility of  $\Delta M$  are 1, 3.

Since,

$$L \subset M, \Delta L = 3$$

Then,

$$\Delta M \neq 1$$

Then,

$$\Delta M = 3$$

Then,

$$\begin{aligned}[M : L] &= \frac{\Delta L}{\Delta M} \\ &= \frac{3}{\Delta M} \\ &= \frac{3}{3} \\ &= 1\end{aligned}$$

Hence, the required index  $[M : L]$  is  $\boxed{1}$ .

2. a

To prove that  $[M : A] = |\det A|$ ,

Suppose  $L, M$  be the lattices with  $L \subset M$  and  $B$  and  $C$  be the basis of lattices  $L$  and  $M$  respectively such that,

$$B = CA$$

Where,  $A$  is an integer matrix.

Suppose  $M$  and  $N$  are the integer lattices in the plane  $\mathbb{R}^2$  and  $L \subset M$ , then

$$[M : L] = \frac{\Delta L}{\Delta M}$$

Since,  $\Delta L$  is the area of the parallelogram of the linear combination of the lattice basis vectors of  $B$  and the area of any matrix is the absolute value of determinant of the matrix.

Then,

$$\Pi(B) = |\det(B)|$$

Since,

$$B = CA$$

Then,

$$\begin{aligned} \Pi(B) &= |\det(B)| \\ &= |\det(CA)| \\ &= |\det(C)| |\det(A)| \end{aligned}$$

Since,  $\Delta M$  is the area of the parallelogram of the linear combination of the lattice basis vectors of  $C$  and the area of any matrix is the absolute value of determinant of the matrix.

Then,

$$\begin{aligned} \Delta M &= \Pi'(C) \\ &= |\det(C)| \end{aligned}$$

Then,

$$\begin{aligned} [M : L] &= \frac{\Delta L}{\Delta M} \\ &= \frac{|\det(B)|}{|\det(C)|} \\ &= \frac{|\det(C)| |\det(A)|}{|\det(C)|} \\ &= |\det(A)| \end{aligned}$$

Hence, it is proved that index  $[M : A] = |\det A|$ .

## Miscellaneous Problem

1. a

Consider the provided statement to describe that the subrings  $S$  of  $\mathbb{C}$  are lattices in the complex number.

[Comment](#)

Step 2 of 2 ^

As it is provided that  $S$  is a subring of  $\mathbb{C}$  such that  $1 \in S$ . Since,  $S$  is a lattice and it is assumed that  $S = \alpha \cdot \mathbb{Z} + \beta \cdot \mathbb{Z}$  where  $\alpha, \beta \in \mathbb{C}$ . Let  $x\alpha + y\beta = 1$  then  $zx\alpha + zy\beta = z \in S$  for any integer  $z$  therefore  $\mathbb{Z} \subset S$ .

Furthermore,  $\beta$  can be also written as  $\frac{1}{y} - \frac{x}{y}\alpha$ , then it is assumed that  $y > 1, \beta^k \in S$  for any integer  $k$  and consequently  $\frac{1+\gamma}{y^k} \in S$  where  $\gamma \in \mathbb{Z}[\alpha]$ .

It means that  $S$  is not discrete even though  $S$  is a lattice is **proved**.

## 2. a

(a)

To prove that one of the ellipse  $x^2 + 5y^2 = p$  or  $x^2 + 5y^2 = 2p$  contains an integer point,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{d}$  and  $\delta^2 = -5$ .

Since,

$$(p) = \overline{P}P$$

Where,  $p$  is a prime.

And the equations of ellipses are,

$$x^2 + 5y^2 = p, x^2 + 5y^2 = 2p$$

First, find the integer solution for the equation of the ellipse,

$$x^2 + 5y^2 = p$$

Suppose  $(0, a)$  be the integer solution of the above equation.

Then,

$$x^2 + 5y^2 = p$$

$$0^2 + 5a^2 = p$$

$$a^2 = \frac{p}{5}$$

Since, the number  $p$  is a prime.

Then, for the integer solution  $5$  must divide  $p$ .

Then, the only possibility for the integer solution is that,

$$p = 5$$

Then,

$$\begin{aligned} a^2 &= \frac{p}{5} \\ &= \frac{5}{5} \\ &= 1 \end{aligned}$$

Then,

$$a = \pm 1$$

Therefore, the integer solutions of the equation of ellipse  $x^2 + 5y^2 = p$  are  $(0, 1), (0, -1)$ .



Now, check whether the points  $(0,1), (0,-1)$  are solution of the equation of ellipse  $x^2 + 5y^2 = 2p$  or not,

Since, the equation of ellipse is,

$$x^2 + 5y^2 = 2p$$

Then, for  $p = 5$ , the points  $(0,1)$  is

$$x^2 + 5y^2 = 2 \cdot 5$$

$$0^2 + 5(1)^2 = 10$$

$$5 = 10$$

Since, above is not possible.

Then, the point  $(0,1)$  is not the solution of the equation of the ellipse  $x^2 + 5y^2 = 2p$ .

In the similar way, the point  $(0,-1)$  is also not a solution of the equation of the ellipse

$$x^2 + 5y^2 = 2p.$$

Then, whenever the equation of ellipse  $x^2 + 5y^2 = p$  has integer solution then the other equation of ellipse  $x^2 + 5y^2 = 2p$  does not has.

In the similar manner, whenever the equation of ellipse  $x^2 + 5y^2 = 2p$  has integer solution then the other equation of ellipse  $x^2 + 5y^2 = p$  does not have.

Hence, **it is proved that exactly one of the ellipses  $x^2 + 5y^2 = p$  or  $x^2 + 5y^2 = 2p$  contains an integer point.**

(b)

To find the property that determines which ellipse has an integer solution,

Suppose  $R = \mathbb{Z}[\delta]$  be the ring of integers in the imaginary quadratic field  $\mathbb{Q}[\delta]$ .

Where,  $\delta = \sqrt{-5}$  and  $\delta^2 = -5$ .

And the equations of ellipses are,

$$x^2 + 5y^2 = p, x^2 + 5y^2 = 2p$$

Suppose  $(x, y) = (m, n)$  be the solution of the equation of ellipse,

$$x^2 + 5y^2 = p$$

Then,

$$m^2 + 5n^2 = p$$

Then,  $(x, y) = (m, n)$  is a solution of the equation of the ellipse  $x^2 + 5y^2 = p$  if and only if,

$$p \equiv m^2 \pmod{5}$$

This implies that, above is hold if and only if,

$$p = 5 \text{ or } p \equiv 1 \pmod{5} \text{ or } p \equiv 4 \pmod{5}$$

Then, in modulo 4

$$\begin{aligned} p &= m^2 + 5n^2 \\ &\equiv m^2 + n^2 \pmod{4} \end{aligned}$$

Then,

$$p \equiv 1 \pmod{4}$$

Then, the combine the condition for  $(m, n)$  to be the solution of the equation of the ellipse

$$x^2 + 5y^2 = p \text{ is,}$$

$$p = 5 \text{ or } p \equiv 1 \pmod{20} \text{ or } p \equiv 9 \pmod{20}$$

This is the necessary and sufficient condition for equation of the ellipse  $x^2 + 5y^2 = p$  to have the integer solution.

Then, equation of the ellipse  $x^2 + 5y^2 = p$  has the integer solution if and only if

$$p = 5 \text{ or } p \equiv 1 \pmod{20} \text{ or } p \equiv 9 \pmod{20}.$$

Hence, the required property that determines the ellipse  $x^2 + 5y^2 = p$  has an integer solution is

that  $\boxed{p = 5 \text{ or } p \equiv 1 \pmod{20} \text{ or } p \equiv 9 \pmod{20}}$

Now, for the integer solution of the equation of ellipse  $x^2 + 5y^2 = 2p$ ,

Since, the class group of the imaginary number field  $\mathbb{Q}[\sqrt{-5}]$  is 2.

Suppose  $q_2$  be the prime ideal over the ramified prime 2.

Then,  $q_2$  is a unique prime ideal and

$$(2) = q_2^2$$

Then,

$$\overline{q_2} = q_2$$

And

$$q_2 = (2, 1 + \sqrt{-5})$$

Then, the condition to have an integer solution for the equation of ellipse  $x^2 + 5y^2 = 2p$  is,

$$\begin{aligned} (2p) &= q_2^2 \overline{P} P \\ &= (\overline{P} q_2)(P q_2) \\ &= (\overline{P} q_2)(P q_2) \end{aligned}$$

Where,  $q_2$  is a unique prime ideal.

Hence, the required property that determines the ellipse  $x^2 + 5y^2 = 2p$  has an integer solution is

that  $\boxed{(2p) = (\overline{P} q_2)(P q_2)}$ .

### 3. a

(a)

To describe the prime ideals in the polynomial ring  $\mathbb{C}[x, y]$  in two variables,

Suppose  $\mathbb{C}[x, y]$  be the polynomial ring with complex coefficients in the variables  $x$  and  $y$ ,

Since, an ideal  $P \neq \mathbb{C}[x, y]$  is called prime ideal in  $\mathbb{C}[x, y]$  if  $AB \subset P$ , then  $A \subset P$  or  $B \subset P$ .

Where,  $A$  and  $B$  are ideals of  $\mathbb{C}[x, y]$ .

And the quotient ring  $\frac{R}{P}$  is integral domain if and only if the ideal  $P$  is prime ideal.

Then,

$$\frac{\mathbb{C}[x, y]}{P} = \frac{\mathbb{C}[x, y]}{(y^2 - x^3 - x - 1)}$$

Since, the polynomial  $(y^2 - x^3 - x - 1)$  is irreducible in  $\mathbb{C}[x, y]$ .

Then, the factor ring  $\frac{\mathbb{C}[x, y]}{(y^2 - x^3 - x - 1)}$  is a field.

Since, every field is an integral domain.

Then, the factor ring  $\frac{\mathbb{C}[x, y]}{(y^2 - x^3 - x - 1)}$  is an integral domain.

Since, the factor ring  $\frac{R}{P}$  is integral domain if and only if the ideal  $P$  is prime ideal.

This implies that, the ideal  $(y^2 - x^3 - x - 1)$  is a prime ideal in the polynomial ring  $\mathbb{C}[x, y]$  with complex coefficients in the variables  $x$  and  $y$ .

In the similar way,  $\frac{\mathbb{C}[x, y]}{(x^2 + y^2 - 1)}$  and  $\frac{\mathbb{C}[x, y]}{(x)(x^2 + y^2 - 1)}$  are integral domains.

This implies that, the ideals  $(x^2 + y^2 - 1)$  and  $(x)(x^2 + y^2 - 1)$  are prime ideals in the polynomial ring  $\mathbb{C}[x, y]$  with complex coefficients in the variables  $x$  and  $y$ .

Then, the entire ideal  $P$  in two variables  $x$  and  $y$  for which the factor ring  $\frac{\mathbb{C}[x, y]}{P}$  is an integral domain is the prime ideal.

Hence, the required prime ideals in the polynomial ring  $\mathbb{C}[x, y]$  in two variables are

$$\boxed{(y^2 - x^3 - x - 1), (x^2 + y^2 - 1), (x)(x^2 + y^2 - 1)}.$$

(b)

To describe the prime ideals in the polynomial ring  $\mathbb{Z}[x]$ ,

Suppose  $\mathbb{Z}[x]$  be the polynomial ring with integer coefficients in the variable  $x$ ,

Since, the ideal  $P = (x)$  is an ideal in the ring of polynomial  $\mathbb{Z}[x]$ ,

Then, the ring  $\frac{\mathbb{Z}[x]}{(x)}$  is a factor ring.

Since,

$$\frac{\mathbb{Z}[x]}{(x)} \cong \mathbb{Z}$$

This implies that, factor ring  $\frac{\mathbb{Z}[x]}{(x)}$  is an integral domain.

Since, the factor ring  $\frac{R}{P}$  is integral domain if and only if the ideal  $P$  is prime ideal.

Then, the ideal  $P = (x)$  is a prime ideal in the polynomial ring  $\mathbb{Z}[x]$  with integer coefficients in the variable  $x$ .

In the similar way, the ideal  $P = (p, x)$  is an ideal in the ring of polynomial  $\mathbb{Z}[x]$ .

Then, the ring  $\frac{\mathbb{Z}[x]}{(p, x)}$  is a factor ring.

Since,

$$\frac{\mathbb{Z}[x]}{(p, x)} = \frac{\mathbb{Z}}{(p)} \cong \mathbb{Z}_p$$

Then, the factor ring  $\frac{\mathbb{Z}[x]}{(p, x)}$  is a field and every field is an integral domain.

Then, the factor ring  $\frac{\mathbb{Z}[x]}{(p, x)}$  is an integral domain.

This implies that, the ideal  $(p, x)$  is a prime ideal.

Hence, the required prime ideals in the polynomial ring  $\mathbb{Z}[x]$  are  $\boxed{(x), (p, x)}$ .

4. a

(a)

To prove Pick's theorem  $\Delta(P) = a + \frac{b}{2} - 1$

Suppose  $L$  be the integer lattice  $\mathbb{Z}^2$  in the plane  $\mathbb{R}^2$  and  $P$  be the polygon in the plane such that vertices are the points of  $L$ .

Since,  $a$  is the number of points of  $L$  in the interior of  $P$  and  $b$  be the number of points of  $L$  on the boundary of  $P$ .

And except for two end points of the brim, entire boundary points along the brim in common are conjoin to the interior points which are conjoin to boundary points because the polygon  $P$  and the lattice  $L$  share a brim.

Suppose  $m$  be the number of points in common.

Then, the numbers of interior points on  $PL$  are,

$$a_{PL} = (a_P + a_L) + (c - 2)$$

And the number of points on the boundary of  $PL$  are,

$$b_{PL} = (b_P + b_L) - 2(c - 2) - 2$$

But,  $a_P$  and  $a_L$  are the number of points of  $P$  and  $L$  respectively.

Then,

$$(a_P + a_L) = a_{PL} - (c - 2)$$

And  $b_P$  and  $b_L$  be the number of points on the boundary of  $P$  and on the boundary of  $L$  respectively.

Then,

$$(b_P + b_L) = b_{PL} + 2(c - 2) + 2$$

Then, the area  $\Delta(P)$ ,

$$\begin{aligned} \Delta(P) &= \left(a_P + \frac{b_P}{2} - 1\right) + \left(a_L + \frac{b_L}{2} - 1\right) \\ &= (a_P + a_L) + \frac{(b_P + b_L)}{2} - 2 \\ &= a_{PL} - (c - 2) + \frac{(b_{PL} + 2(c - 2) + 2)}{2} - 2 \\ &= a_{PL} + \frac{b_{PL}}{2} - 1 \end{aligned}$$

Since,  $a_{PL}$  is the number of points of  $L$  in the interior of  $P$ .

Then,

$$a_{PL} = a$$

And  $b_{PL}$  is the number of points of  $L$  on the boundary of  $P$ .

Then,

$$b_{PL} = b$$

Then, the area  $\Delta(P)$ ,

$$\begin{aligned} \Delta(P) &= a_{PL} + \frac{b_{PL}}{2} - 1 \\ &= a + \frac{b}{2} - 1 \end{aligned}$$

Hence, it is proved that Pick's theorem  $\Delta(P) = a + \frac{b}{2} - 1$ .

(b)

To derive the proposition that  $[P:L] = \frac{\Delta L}{\Delta P}$ .

Suppose  $L$  be the integer lattice  $\mathbb{Z}^2$  in the plane  $\mathbb{R}^2$  and  $P$  be the polygon in the plane such that vertices are the points of  $L$ .

Then, the area of the lattice  $L$  is,

$$\Delta(L) = a + \frac{b}{2} - 1$$

Where,  $a$  is the number of points of  $L$  in the interior of  $P$  and  $b$  be the number of points of  $L$  on the boundary of  $P$ .

Suppose  $c$  be the number of points of  $P$  in the interior of the plane and  $d$  be the number of points of  $P$  on the boundary of the plane.

Then, by Pick's theorem, the area of polygon  $P$  in the plane is,

$$\Delta(P) = c + \frac{d}{2} - 1$$

Now,

$$\begin{aligned} \frac{\Delta L}{\Delta P} &= \frac{a + \frac{b}{2} - 1}{c + \frac{d}{2} - 1} \\ &= \frac{2a + b - 2}{2c + d - 2} \end{aligned}$$

Since,  $L \subset P$  and the number of points of  $L$  in the interior of  $P$  is  $a$  and the number of points of  $L$  on the boundary of  $P$  is  $b$ .

Then, the area  $\frac{2a + b - 2}{2c + d - 2}$  represents the additive cosets of  $L$  in  $P$ .

Since, the additive cosets of  $L$  in  $P$  are denoted by  $[P:L]$ .

Then,

$$\begin{aligned} \frac{2a + b - 2}{2c + d - 2} &= [P:L] \\ \frac{\Delta L}{\Delta P} &= [P:L] \end{aligned}$$

Hence, it concludes that  $[P:L] = \frac{\Delta L}{\Delta P}$

## Chapter 14

### Section 1

1. a

Let  $\varphi : V \rightarrow V$  be a homomorphism. By definition 14.1.4, it holds for all  $v \in V$  that  $\varphi(v) = \varphi(v \cdot 1) = v\varphi(1)$ . Therefore, if we denote  $\varphi(1) = a \in R$ , each homomorphism is of form  $\varphi(v) = v \cdot a$ . Now we just need to confirm that all  $\varphi$  of this form are indeed  $R$ -module homomorphisms. Let  $r \in R$  and define  $\varphi(v) = vr$ . Then we have:

$$\begin{aligned}\varphi(v + v') &= (v + v')r = vr + v'r = \varphi(v) + \varphi(v') \\ \varphi(sv) &= (sv)r = s(vr) = s\varphi(v)\end{aligned}$$

Therefore, all homomorphisms  $\varphi$  on an  $R$ -module  $R$  are of form  $\varphi(v) = vr$  for some  $r \in R$ .

#### Result

2 of 2

Here we use the multiplicative criterion to show that all homomorphisms on  $R$  are simply multiplication with a certain element.

2. a

Let  $V$  be a  $\mathbb{Q}$  module. Let us consider, for any  $v$  in  $V$  and  $n \in \mathbb{N}$ ,  $\frac{1}{n}v$ . By 14.1.1, it needs to hold that

$$\underbrace{\frac{1}{n}v + \frac{1}{n}v + \dots + \frac{1}{n}v}_{n \text{ times}} = n \left( \frac{1}{n}v \right) = \left( n \cdot \frac{1}{n} \right) v = v. \text{ Therefore, } w = \frac{1}{n}v \text{ is an element of } V \text{ such that } nw = v.$$

Assume that there exists another such element  $w'$ . Then it holds that:

$$\begin{aligned}nw - nw' &= v - v \\ n(w - w') &= 0 \\ w - w' &= 0 \\ w &= w'.\end{aligned}$$

Therefore,  $\frac{1}{n}v$  is uniquely determined. Now let  $m \in \mathbb{Z}$  and consider  $\frac{m}{n}v$ . Imitating the constructions for the  $\mathbb{Z}$ -module from the book and the construction above, we see that for positive  $m$ ,  $w = \frac{m}{n}v$  is an element such that  $w = m \cdot \frac{1}{n}v$ . Clearly, if  $w' = m \frac{1}{n}v$ , then  $w - w' = 0$  so  $w = w'$ . The argument for  $m = 0$  and  $m < 0$  is very similar and so the claim follows.

#### Result

2 of 2

To obtain the solution to this exercise, we do a construction similar to the proof of  $\mathbb{Z}$ -module structure of an abelian group, although here we verify our claim.

3. a



$R$  is the set  $\{a + b\alpha \mid a, b \in \mathbb{Z}\}$ . Let  $m \in \mathbb{Z}$ . Therefore, a coset  $A_{a_1, b_1}$  in  $R/mR$  for  $a_1, b_1 \in \mathbb{Z}$  is of the form  $\{a_1 + b_1\alpha + ma + mb\alpha \mid a, b \in \mathbb{Z}\} = \{a_1 + ma + (b_1 + mb)\alpha \mid a, b \in \mathbb{Z}\}$ . Notice that if  $a_0 = a_1 + km$  for some  $k \in \mathbb{Z}$ , then  $A_{a_0, a_1} = A_{a_1, b_1}$ . Similarly, if  $b_0 = b_1 + km$  for some  $k \in \mathbb{Z}$ , then  $A_{a_1, b_0} = A_{a_1, b_1}$ . Therefore, we all cosets are equal to a coset  $A_{i, j}$  for  $i, j \in \{0, 1, \dots, m-1\}$ . It is also easily confirmed that  $A_{i, j} \neq A_{k, l}$  if  $i \neq k$  or  $j \neq l$ . This means that there are  $m^2$  cosets, i.e.,  $R/mR$  is finite group of order  $m^2$ .

## Result

2 of 2

The solution to this exercise is easily obtained as soon as we write out what are the cosets in this quotient ring.

## 4. a

(a) Let  $S$  be a simple  $R$ -module and let  $s \in S, s \neq 0$ . Define a function  $\psi : R \rightarrow S, \psi(r) = rs$ . Since  $s$  is non-zero, the image of  $\psi$  is non-zero. Also, the image of  $\psi$  is a submodule, but by definition of a simple module, it cannot be a proper submodule. Therefore,  $\Im(\psi) = S$ . Hence,  $\psi$  is surjective.

By the [First Isomorphism Theorem \(14.1.6 \(c\)\)](#),  $S$  is isomorphic to  $R/\ker \psi$ . Assume that  $\ker \psi$  is not a maximal ideal. Then it is contained in some (maximal, by Proposition A.3.5) ideal  $M'$ . But  $R/M'$  is then isomorphic to a submodule of  $S$  by the [Correspondence Theorem \(14.1.6 \(d\)\)](#). This is a contradiction because  $S$  is simple by assumption. Therefore,  $\ker \psi$  is a maximal ideal.

## Step 2

2 of 3

(b) Note that  $\Im \varphi$  is a submodule of  $S'$ . Since  $S'$  is simple,  $\Im \varphi$  can only be the zero module or the whole  $S'$ . If  $\Im \varphi = 0$ , then  $\varphi$  is a [zero homomorphism](#) and the statement holds. If  $\Im \varphi = S'$ ,  $\varphi$  is [surjective](#). Let us consider  $\ker \varphi$ . It is a submodule of  $S$ , which is simple, so the kernel can only be the zero module or the whole  $S$ . If  $\ker \varphi = 0$ , then  $\varphi$  is injective as well as surjective, so  $\varphi$  is an [isomorphism](#). If  $\ker \varphi = S$ , then  $\Im \varphi = S'$  implies that  $S'$  is the zero module, which is a [contradiction](#) with the definition of a simple module.

## Result

3 of 3

To obtain the solution to this exercise, we use the fact that both the image and the kernel of a module are submodules, as well as different parts of 14.1.6.

# Section 2

## 1. a

Assume that  $M$  is a free  $R$ -module and that

it has a basis  $B$  of at least two elements

. Let  $a, b \in B$ . Since  $a, b \in M$  and  $M \subseteq R$ , it holds that  $a, b \in R$ . Therefore, there exists a linear combination  $b \cdot a + (-a) \cdot b = 0$ . We arrive at a contradiction, since this means that the basis is not linearly independent. Therefore, if there is a basis, it can only have one element. On the other hand, if the basis consists of a single element, then

$M$  is generated by only one element

. We claim, however, that the ideal generated by  $x$  and  $y$  is not a principal ideal. As soon as we prove this, the exercise is solved.

Assume that the ideal generated by  $x$  and  $y$  is principal, that is, it is generated by a single element  $z$ . For that to be the case, it needs to hold that for each  $p_1, p_2 \in R$  there exists a  $p \in R$  such that:  $p_1x + p_2y = pz$ . If we consider  $p_1 = p_2 = 1$ , we see that  $z$  can be of degree at most 1. If  $\deg z = 0$ , then the ideal generated by  $z$  is the whole  $R$ , which is not the case for  $M$ . Therefore,  $\deg z = 1$ , so  $z = ax + by + c$ . Now  $p_1 = 0, p_2 = 1$  implies that  $a = c = 0$ , while  $p_1 = 1, p_2 = 0$  implies that  $b = c = 0$ . Therefore,  $z$  must be of degree zero, which is a contradiction with the above claim. Hence,

$M$  is not a free module

## Result

3 of 3

We solve this exercise via contradiction and case analysis of a basis of 2 or more elements versus a single element basis.

## 2. a

Zero ring  $Z$  clearly has the property that every finitely generated  $Z$ -module is free. Therefore, assume  $R$  has the aforementioned property and is not a zero ring. Let  $I$  be a proper non-zero ideal in  $R$ . Then  $R/I$  is an  $R$ -module which is finitely generated (principal!) by  $1 + I$ . By assumption,  $R/I$  is free with some basis  $B$ . Let  $b \in B$  and let  $i \in I, i \neq 0$ . Such an  $i$  exists because  $I$  is non-zero by assumption. It holds that:  $i(b + I) = I$ , which means that the basis is linearly dependent because  $I$  is the identity in  $R/I$ . This is a contradiction, so there does not exist a proper non-zero ideal in  $R$ . Thus,  $R$  is a field.

## Result

2 of 2

We prove this claim using the properties of ideals.

## 3. a

Consider the provided statement to prove the provided condition. As it is given that  $A$  be the matrix of a homomorphism such that  $\varphi: \mathbb{Z}^n \rightarrow \mathbb{Z}^m$ .

(a)

Let  $a_1, a_2, \dots, a_n$  and  $b_1, b_2, \dots, b_m$  be the bases of the free modulo  $\mathbb{Z}^n$  and  $\mathbb{Z}^m$  respectively, such that

$$\varphi(a_1, a_2, \dots, a_n) = (b_1, b_2, \dots, b_m)A$$

It is clear that if  $\varphi$  is injective if and only if the homogeneous linear equations  $AX = 0$  only have zero solution in the rational number field  $\mathbb{Q}$ . Therefore, it can be said that the rank of  $A$  as a real matrix is  $n$ . Hence, provided statement is **proved**.

(b)

It is assumed that the GCD of the determinants of the  $m \times m$  minors of  $A$  is  $k > 1$ . From the definition of the determinant of a matrix, it can be seen that  $\text{Im } \varphi$  must be contained in  $k\mathbb{Z}^m$ . Therefore,  $\varphi$  is not surjective.

It is also assumed that the GCD of the determinants of the  $m \times m$  minors of  $A$  is 1 and this implies that  $n \geq m$ . Then there exists an  $\mathbb{Z}$ -invertible  $n \times n$  matrix  $B$  such that,

$$AB = (A_1, A_2)$$

Where  $A_1$  is an  $\mathbb{Z}$ -invertible  $m \times m$  matrix and  $A_2$  is an  $m \times (n-m)$  matrix. Let  $\phi$  be the auto-morphism of  $\mathbb{Z}^n$  corresponding to  $B$ . Therefore,

$$\begin{aligned} (b_1, \dots, b_m)(A_1, A_2) &= (b_1, \dots, b_m)AB \\ &= \varphi(a_1, \dots, a_n)B \\ &= \phi\varphi(a_1, \dots, a_n) \end{aligned}$$

Since  $A_1$  is invertible, it can be seen that  $\phi\varphi$  is surjective. Moreover, since  $B$  is invertible so it can be said that  $\varphi$  is surjective. Hence, provided statement is **proved**.

4. a

Consider the provided statement to prove the given condition as it is provided that  $I$  be an ideal of a ring  $R$ .

(a)

The circumstances under which  $I$  is a free  $R$ -module if and only if it is a principal ideal which is generated by an element  $\alpha$  and that is not a zero divisor in  $R$ .

This statement may or may not contradict in the case of  $\alpha = 0$ . It is assumed that,

$$\begin{aligned} I &= (\alpha) \\ &= \alpha R \end{aligned}$$

Then it is claimed that, a basis for  $I$  is the set  $\{\alpha\}$ . So  $\{\alpha\}$  is linearly independent if  $r\alpha = 0, r \neq 0$  then  $\alpha$  will be a zero divisor and  $\{\alpha\}$  is a spanning set because from the definition,

$$I = R\alpha$$

Conversely, it is assumed that  $I \subset R$  is a free  $R$ -module and let  $\alpha_1, \dots, \alpha_n \in I$  be a basis. If  $n \geq 2$  then it is observed that the following given relation is linear dependence relation.

$$\alpha_2 \times \alpha_1 + (-\alpha_1) \times \alpha_2$$

Therefore, if  $I$  is a free  $R$ -module then  $n = 1$  and  $I = (\alpha_1) = \alpha$ . Furthermore, if  $\alpha$  is a zero divisor then  $\alpha\beta = 0, \beta \neq 0$  then  $I$  does not have a basis that is a zero-element set clearly doesn't span and if  $r\alpha$  is in the basis for some  $r \in R$  then  $\beta(r\alpha) = r(\alpha\beta) = 0$  is a non-trivial linear dependence relation.

Therefore,  $I$  is a free  $R$ -module if and only if it is a principal ideal which is generated by an element  $\alpha$ .

(b)

It is assumed that every finitely generated  $R$ -module is free and as it is given that  $I$  be a proper ideal of  $R$  then  $\frac{R}{I}$  is a finitely generated  $R$ -module, if  $I$  is not the zero ideal then  $\frac{R}{I}$  is not isomorphic to  $R^n$  for any  $n$  that is a zero-element does not span and any element is linearly dependent.

Therefore, if every finitely generated  $R$ -module is free then  $R$  does not have proper non-zero ideals.

## Section 3

1. a

Consider the provided statement to prove that if  $f$  is the zero function then  $f$  is the zero polynomial. As it is provided that  $\bar{f}$  is the function that defined by evaluation of a complex polynomials  $f(x_1, \dots, x_n)$ .

[Comment](#)

---

Step 2 of 2 ^

As it is assumed that  $f(x_i) \in \mathbb{Z}[x_1, \dots, x_n]$  be a polynomial identity and whose corresponding real polynomial function is  $\bar{f}: R^n \rightarrow R$

From the proposition 3.8, if  $\bar{f}$  is identically zero then  $f = 0$ , this is property of real numbers and rest of the proof is given for the complex numbers is identical.

Therefore the image of  $f$  under the canonical ring homomorphism  $\mathbb{Z}[x_i] \rightarrow R[x_i]$  is zero.

Hence, provided statement is **proved**.

2. a

An identity is an equality relation  $A = B$ , such that  $A$  and  $B$  produce the same value as each other regardless of the values substituted for the variables.

[Comment](#)

---

Step 2 of 3 ^

To show: that it might be convenient to verify an identity only for the real numbers

For this, consider  $R$  is a set of real numbers which is a ring with identity  $I_R$

Let  $S$  be the subring of  $R$  and that;

$$I_R \in S$$

So, consider  $I_R$  to be the identity element in  $S$

This,  $S$  is a ring with identity  $I_s$ , that is;

$$I_s = I_R$$



Now, suppose that  $x$  is a unit in  $S$

Then, there exists an element;

$$y \in S$$

Such that;

$$xy = I_S$$

Hence;

$$xy = I_R$$

Since;

$$S \subseteq R; x, y \in R$$

The equation;

$$xy = I_R$$

This implies that  $x$  is a unit in  $R$

**Therefore, it is convenient to verify an identity only for the real numbers**

3. a

Consider the provided statement to prove that the trace of the linear operator on the space is the product  $(\text{trace } A)(\text{trace } B)$  by using permanence of identities.

[Comment](#)

Step 2 of 3 ^

As it is provided that the order of matrices  $A$  and  $B$  are  $m \times m$  and  $n \times n$  respectively then let  $X_A$  be the left multiplication by  $A$  and  $Y_B$  be the right multiplication by  $B$ . Then, it can split  $R^{m \times n}$  into the direct sum of  $X_A$ -invariant  $n$  copies of  $R^n$  to the provided matrix the columns of  $M$ .

$X_A$  Acts by the usual multiplication of a matrix by a column vector, therefore

$$\det(X_A) = [\det(A)]^n$$

This process is similar as,

$$\det(Y_B) = [\det(B)]^m$$

Therefore,

$$\begin{aligned} \det(X_A Y_B) &= \det(X_A) \det(Y_B) \\ &= [\det(A)]^n [\det(B)]^m \end{aligned}$$

For compute the  $\text{trace}(X_A Y_B)$  clear the previous argument to  $X_A Y_B = A \otimes B^*$ . It can be think is shown as below.

$$\begin{aligned} R^{m \times n} &= \text{Hom}(R^m, R^n) \\ &= R^m \otimes (R^n)^* \end{aligned}$$

Hence,  $\text{trace}(X_A Y_B) = \text{trace}(X_A) \text{trace}(Y_B)$  is **proved**.

4. a

(a)

To determine that whether or not permanence of identities allows the result to be carried over from complex numbers to an arbitrary commutative ring,

Suppose  $\mathbb{C}$  be the set of complex numbers and  $A, B, C$  be the compatible matrices with entries from set of complex numbers.

Since, matrices satisfy the associative law under multiplication.

Then, for the matrices  $A, B, C$

$$(AB)C = A(BC)$$

Then, the above can be written in the form of a function  $f$  as,

$$f(A, B, C) = (AB)C - A(BC)$$

Suppose  $X, Y, Z$  be the indeterminate  $n \times n$  matrices.

Then,

$$(XY)Z = X(YZ)$$

Then,

$$f(X, Y, Z) = (XY)Z - X(YZ)$$

Then,  $f(X, Y, Z)$  is the polynomial and an element of the ring.

Then, the associative law of the matrices under multiplication is similar to an identity.

Since, the certain properties of matrices with entries in a field continue to hold when the entries are in the ring only if the property is an identity.

Then, the associative law of the matrices under multiplication holds in the ring.

Hence, **it concludes that permanence of identity allows the result for the associative law of matrix multiplication to be carried over from complex numbers to an arbitrary commutative ring.**

(b)

To determine that whether or not permanence of identities allows the result to be carried over from complex numbers to an arbitrary commutative ring,

Suppose  $\mathbb{C}$  be the set of complex numbers and  $p(t)$  be the characteristic polynomial of a  $n \times n$  matrix  $A$ .

Then, by the Cayley-Hamilton theorem, the characteristic polynomial  $p(t)$  is,

$$\begin{aligned} p(t) &= \det(tI - A) \\ &= t^n - (\text{trace } A)t^{n-1} + (\text{indeterminate terms}) + (-1)^n (\det A) \end{aligned}$$

Suppose  $X$  be the indeterminate  $n \times n$  matrix.

Then,

$$\begin{aligned} p(t) &= \det(tI - X) \\ &= t^n - (\text{trace } X)t^{n-1} + (\text{indeterminate terms}) + (-1)^n (\det X) \end{aligned}$$

Then, the polynomial  $p(t)$  is the polynomial in  $2n^2$  complex variables and this is an element of the ring.

Hence, **it concludes that permanence of identity allows the result for the Cayley-Hamilton theorem to be carried over from complex numbers to an arbitrary commutative ring.**



(c)

To determine that whether or not permanence of identities allows the result to be carried over from complex numbers to an arbitrary commutative ring,

Suppose  $\mathbb{C}$  be the set of complex numbers and a system of  $n$  linear equations for  $n$  unknowns such that,

$$Ax = B$$

Where,  $A$  is the non-singular coefficient matrix, the column vector  $x = (x_1, x_2, \dots, x_n)^T$  and  $B$  is the constant matrix.

Then, by the Cramer's rule,

$$x_i = \frac{\det(A_i)}{\det(A)}$$

Where,  $A_i$  is the matrix formed by replacing the  $i$ -th column of the coefficient matrix  $A$  and  $i = 1, 2, \dots, n$ .

Suppose  $X$  be the indeterminate  $n \times n$  matrix.

Then,

$$x_i = \frac{\det(X_i)}{\det(X)}$$

Where,  $X_i$  is the matrix formed by replacing the  $i$ -th column of the matrix  $X$  and  $i = 1, 2, \dots, n$ .

Since, both  $\det(X)$  and  $\det(X_i)$  are polynomials in a unknown.

Then,  $\frac{\det(X_i)}{\det(X)}$  need not be the polynomial.

This implies that,  $\frac{\det(X_i)}{\det(X)}$  need not be an element of the ring.

Hence, **it concludes that permanence of identity does not allows the result for Cramer's rule to be carried over from complex numbers to an arbitrary commutative ring.**

(d)

To determine that whether or not permanence of identities allows the result to be carried over from complex numbers to an arbitrary commutative ring,

Suppose  $\mathbb{C}$  be the set of complex numbers.

Then, any product, quotient of complex numbers is again a complex number.

And product of polynomials is again a polynomial.

And the differentiation of polynomial is again a polynomial.

---

[Comment](#)

---

Step 5 of 7 ^

But, the quotient of two polynomial need not be a polynomial.

Then, the quotient of two polynomials need not be an element of the ring.

Hence, **it concludes that permanence of identity does not allows the result for product rule, quotient rule, and chain rule for the differentiation of the polynomials to an arbitrary commutative ring.**

(e)

To determine that whether or not permanence of identities allows the result to be carried over from complex numbers to an arbitrary commutative ring,

Suppose  $\mathbb{C}$  be the set of complex numbers and  $f(x)$  be a polynomial in indeterminate  $x$  of degree  $n$  in.

Then, the polynomial  $f(x)$  can have at most  $n$  roots in any field and this is an identity.

Since, the certain properties of polynomials with coefficients from the field continue to hold when the coefficients are in the ring only if the property is an identity.

Then, the polynomial  $f(x)$  can have at most  $n$  roots in any ring.

Hence, **it concludes that permanence of identity allows the result for the fact that a polynomial of degree  $n$  has at most  $n$  roots in any arbitrary ring.**

(f)

To determine that whether or not permanence of identities allows the result to be carried over from complex numbers to an arbitrary commutative ring,

Suppose  $\mathbb{C}$  be the set of complex numbers and  $g(x)$  be a complex valued function that is infinitely differentiable at any number  $c$ .

Then, the Taylor's series is,

$$g(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n$$

Then, Taylor's expansion of the polynomials  $g(x)$  is,

$$g(x) = g(b) + g'(b)(x-b) + \frac{g''(b)}{2!}(x-b)^2 + \dots$$

Then, the Taylor's expansion of the polynomials  $g(x)$  is similar to an identity.

Since, the certain properties of polynomials with coefficients from the field continue to hold when the coefficients are in the ring only if the property is an identity.

Then, Taylor's expansion of a polynomial is an element of the ring.

Hence, **it concludes that permanence of identity allows the result for Taylor's expansion of a polynomial in any arbitrary ring.**

## Section 4

1. a

(a)

Here the objective is to transform the matrix in diagonal form and the matrix are,

$$\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix} \text{ and } \begin{bmatrix} 3 & 1 & -4 \\ 2 & -3 & 1 \\ -4 & 6 & -2 \end{bmatrix}.$$

Consider the provided statement to reduce matrix to diagonal form by using integer row and column operations.

Provided matrix is,  $\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$

**Step1:** Interchange  $C_1 \leftrightarrow C_2$  is shown as below,

$$\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$$

**Step2:** Perform  $C_2 \rightarrow C_2 - 3 \cdot C_1$  is shown as below,

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3-3 \cdot 1 \\ 2 & -1-2 \cdot 3 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 \\ 2 & -7 \end{bmatrix}$$

**Step3:** Perform  $R_2 \rightarrow R_2 - 2 \cdot R_1$  is shown as below,

$$\begin{bmatrix} 1 & 0 \\ 2 & -7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 2-2 & -7-0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 \\ -1 & -7 \end{bmatrix}$$

**Step4:** Perform  $R_2 \rightarrow R_2 + R_1$  is shown as below,

$$\begin{bmatrix} 1 & 0 \\ -1 & -7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ -1+1 & -7+0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -7 \end{bmatrix}$$

Hence, the diagonal form of the matrix is  $\begin{bmatrix} 1 & 0 \\ 0 & -7 \end{bmatrix}$ .

Provided matrix is,  $\begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix}$

**Step1:** Perform  $R_2 \rightarrow \frac{R_2}{2}$  then,

$$\begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

**Step2:** Perform  $C_3 \rightarrow 2C_3 - C_1$  then,

$$\begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 7 & 2 \cdot 2 - 4 \\ 1 & 2 & 3 \cdot 2 - 1 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 4 & 7 & 0 \\ 1 & 2 & 5 \end{bmatrix}$$

**Step3:** Perform  $R_2 \rightarrow 4R_2 - R_1$  then,

$$\begin{bmatrix} 4 & 7 & 0 \\ 1 & 2 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 7 & 0 \\ 4-4 & 8-7 & 20-0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 4 & 7 & 0 \\ 0 & 1 & 20 \end{bmatrix}$$

**Step4:** Perform  $R_1 \rightarrow 7R_2 - R_1$  then,

$$\begin{bmatrix} 4 & 7 & 0 \\ 0 & 1 & 20 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 7-7 & -140 \\ 0 & 1 & 20 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 4 & 0 & -140 \\ 0 & 1 & 20 \end{bmatrix}$$

**Step5:** Perform  $C_3 \rightarrow \frac{C_3}{20}$  then,

$$\begin{bmatrix} 4 & 0 & -140 \\ 0 & 1 & 20 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 0 & -7 \\ 0 & 1 & 1 \end{bmatrix}$$

**Step6:** Perform  $C_3 \rightarrow C_3 - C_2$  then,

$$\begin{bmatrix} 4 & 0 & -7 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 0 & -7 \\ 0 & 1 & 1-1 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 4 & 0 & -7 \\ 0 & 1 & 0 \end{bmatrix}$$

**Step7:** Perform  $C_1 \rightarrow 2C_1 + C_3$  then,

$$\begin{bmatrix} 4 & 0 & -7 \\ 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 8-7 & 0 & -7 \\ 0 & 1 & 0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 0 \end{bmatrix}$$

**Step9:** Perform  $C_3 \rightarrow 7C_1 + C_3$  then,

$$\begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -7+7 \\ 0 & 1 & 0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Hence, the diagonal form of the matrix is  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ .

Provided matrix is,  $\begin{bmatrix} 3 & 1 & -4 \\ 2 & -3 & 1 \\ -4 & 6 & -2 \end{bmatrix}$

Now by using MATLAB, the diagonalizing is shown as below,

**Step1:** Enter the matrix in MATLAB is as below;

```
>> A=[3 1 -4;
2 -3 1;
-4 6 -2]
```

A =

```
3    1   -4
2   -3    1
-4    6   -2
```

**Step2:** The Eigen value (EVe) and Eigen vector (EVa) is calculated as below;

```
>> [EVe,EVa]=eig(A)
```

EVe =

```
-0.9102   -0.3789    0.5774
-0.1852    0.4139    0.5774
0.3704   -0.8277    0.5774
```

EVa =

```
4.8310     0     0
0   -6.8310     0
0     0    0.0000
```

**Step3:** The diagonalize matrix is shown as below;

```
>> C=inv(EVe)*A*EVe
```

C =

```
4.8310    -0.0000    0.0000
0.0000   -6.8310    0.0000
0.0000   -0.0000    0.0000
```

Hence, the diagonal form of the matrix is  $\begin{bmatrix} 4.83 & 0 & 0 \\ 0 & -6.83 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ .

**(b)**

To find the bases of  $V$  and  $L$  as it is provided that  $V = \mathbb{Z}^2$  and  $L = AV$ .

Now drawing the lattice spanned by the vectors  $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$  which is left in the notebook. The commensurate bases are  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  for  $\mathbb{Z}^2$  and  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \end{bmatrix}$  for  $L$ , as this is easily too established from the picture.

**(c)**

Here the objective is to transform the matrix in diagonal form.

The integer matrices  $Q^{-1}$  and  $P$  by diagonalize the provided matrix is shown as below;

The provided matrix is,  $A = \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix}$

The provided matrix is diagonalized is as follow by using several row and column operations,

$$\begin{aligned} \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix} &= \begin{bmatrix} 4 & -1 & 2 \\ 2 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 4 & 2 \\ 0 & 6 & -2 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Hence, the value of  $Q^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$  and  $P = \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$ .

2. a

Consider the provided statement to prove that  $d_1$  is the greatest common divisor of the entries  $a_{ij}$  of matrix  $A$ .

[Comment](#)

Step 2 of 2 ^

Let  $M$  be a minor of the provided matrix  $A$  is shown as below;

$$M = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & d_n & \cdots & 0 \\ & & & \cdots & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

Then  $d_1 d_2 \cdots d_i$  is the greatest common divisor of any  $i \times i$  minor of  $M$ . Since  $d_1 | d_2 | \cdots | d_n$  is a  $i \times i$  minor which is zero or the product of  $i$  distinct  $d_j$  will be divisible by the product of the first  $i$   $d_j$  and this product is the greatest because  $d_1 \cdots d_i$  is a possible  $i \times i$  minor.

From the theorem 14.4.6, any integer matrix can be diagonalized into a form like  $M$  by invertible row and column operations. Therefore, if  $N$  is a matrix where  $d_1 \cdots d_i$  is the GCD of all the  $i \times i$  minors.

Hence,  $d_1$  is the greatest common divisor of the entries  $a_{ij}$  is **proved**.

3. a

Consider the provided statement to determine all integer solutions to the system of equations

and also find a basis for the space. The matrix is,  $A = \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix}$

[Comment](#)

Step 2 of 3 ^

The provided matrix is diagonalized is as follow by using several row and column operations,

$$\begin{aligned} \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix} &= \begin{bmatrix} 4 & -1 & 2 \\ 2 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 4 & 2 \\ 0 & 6 & -2 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$



Therefore, let  $X = \begin{bmatrix} x & y & z \end{bmatrix}^T$  then  $AX = 0$  means,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 4 & 7 & 2 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 4x+7y+2z \\ x+2y+3z \\ z \end{bmatrix} = 0$$

As the first matrix is diagonal, then it is concluded that there are no restrictions on the variable  $z$ .

$$4x + 7y + 2z = 0 \dots\dots (1)$$

$$x + 2y + 3z = 0 \dots\dots (2)$$

From equation (2), the value of  $x$  is as  $x = -2y - 3z$  then, substitute this value into equation (1),

$$4(-2y - 3z) + 7y + 2z = 0$$

$$-8y - 12z + 7y + 2z = 0$$

$$-y - 10z = 0$$

$$y = -10z$$

After putting the value of  $y = -10z$  then the value of  $x$  is calculated as below,

$$x = -2y - 3z$$

$$= 20z - 3z$$

$$x = 17z$$

Therefore, the solutions are of the form of  $\begin{bmatrix} 17z \\ -10z \\ z \end{bmatrix}$  integer, if  $z \in \mathbb{Z}$ .

#### 4. a

Consider the provided statement to find a basis for the  $\mathbb{Z}$ -module of the system of equations.

Provided system equations are,

$$x + 2y + 3z = 0 \dots\dots (1)$$

$$x + 4y + 9z = 0 \dots\dots (2)$$

It is easier to solve manually to formally diagonalize, therefore from equation (1)

$$x + 2y + 3z = 0$$

$$x = -(2y + 3z) \dots\dots (3)$$

Now put the value of  $x$  into equation (2), then

$$x + 4y + 9z = 0$$

$$-2y - 3z + 4y + 9z = 0$$

$$2y + 6z = 0$$

$$2(y + 3z) = 0$$

$$y + 3z = 0$$

$$y = -3z$$

Then from equation (3) when the value of  $y$  is substituting,

$$x = -2y - 3z$$

$$= -2(-3z) - 3z$$

$$= 6z - 3z$$

$$= 3z$$

Therefore, from the above calculation it is observed that, the solution have the form as  $z \begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix}$

and the basis is  $\left\{ \begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix} \right\}$ .

#### 5. a

Complex numbers are defined as those which can be written in the form of;

$$a + ib$$

Where;

$$a, b = \text{integers}$$

$$i = \sqrt{-1}$$

[Comment](#)

#### Step 2 of 5 ^

Consider the complex numbers  $\alpha, \beta, \gamma$  are linearly independent over the rationals

Then;

$$(l\alpha - [l\alpha], m\beta - [m\beta], n\gamma - [n\gamma])$$

For  $l, m, n = 1, 2, 3, 4, \dots$

Let these be dense in the unit square.

Further consider an example of a ring  $R = \mathbb{Z}[\sqrt{-2}]$

It is a lattice in the complex plane. It has the points with integer coordinates having basis  $1, \sqrt{2}i$

This means that the mesh of the lattice is a rectangle

In this case, the base is parallel to the real axis that has a length of 1 and a height of length  $\sqrt{2}$

[Comment](#)

#### Step 4 of 5 ^

Now, consider another example;

$$R = \mathbb{Z}\left[\frac{1 + \sqrt{-3}}{2}\right]$$

Clearly,  $\frac{1 + \sqrt{-3}}{2}$  is a cube root of -1. That is the fundamental parallelogram in the rhombus

which is the union of two equilateral triangles

So, in this case a triangular lattice is formed. But if both the triangles are set together then again the mesh becomes a rectangle as given below;

DIAGRAM

Thus from the above explanation it can be said that any complex number of the form of  $\alpha, \beta$  and  $\gamma$  adjoined with the integer  $l, m, n$ , that is;

$$(l\alpha - [l\alpha], m\beta - [m\beta], n\gamma - [n\gamma])$$

That means it should have a mesh of the form of a rectangle.

**Thus, the set of integer with linear combinations  $\{l\alpha + m\beta + n\gamma / l, m, n \in \mathbb{Z}\}$  forms a lattice in complex plane if there mesh forms a rectangle in shape**

6. a

**Given:** Consider a homomorphism given by multiplication by an integer matrix  $A$  given by

$$\varphi : \mathbb{Z}^k \rightarrow \mathbb{Z}^k.$$

**To Prove:** The image of  $\varphi$  is of finite index if and only if  $A$  is non-singular that if so, then the index is equal to  $|\det(A)|$ .

**Proof:** Consider the subgroup  $\text{Im}(\varphi)$  of  $\mathbb{Z}^k$ .

Since  $\mathbb{Z}^k$  is abelian, note that the index of  $\text{Im}(\varphi)$  is given by  $|\mathbb{Z}^k / \text{Im}(\varphi)|$ .

Now recall that, if  $A$  is an integer matrix, then there exist products  $Q$  and  $P$  of elementary integer matrices of appropriate sizes, so that  $A' = Q^{-1}AP$  is diagonal, and if  $d_1, d_2, \dots, d_k$  are diagonal entries of each row down, then  $d_1|d_2|d_3|\dots|d_{k-1}|d_k$ .

So it follows that there exists  $P$  and  $Q$  in  $\text{GL}_k(\mathbb{Z})$  such that  $A' = Q^{-1}AP$  is diagonal with diagonal entries  $d_1|d_2|d_3|\dots|d_{r-1}|d_r$  followed by  $k - r$  zeroes.

And therefore we have

$$\begin{aligned} \mathbb{Z}^k / \text{Im}(\varphi) &= \mathbb{Z}^k / A\mathbb{Z}^k \\ &\sim \mathbb{Z}^k / A'\mathbb{Z}^k \\ &= \prod_{i=1}^r (\mathbb{Z}/d_i\mathbb{Z}) \times \mathbb{Z}^{k-r}. \end{aligned}$$

Therefore notice that  $|\mathbb{Z}^k / \text{Im}(\varphi)|$  is finite if and only if  $r = k$ .

Now this is true if and only if

$$\begin{aligned} \det(A) &= \det(QA'P^{-1}) \\ &= \det(Q)\det(A')\det(P^{-1}) \\ &= \det(Q)\det(P^{-1}) \prod_{i=1}^k d_i \\ &\neq 0 \dots \dots \dots (*) \end{aligned}$$

This follows that  $A$  is non-singular.

Now consider that if  $|\mathbb{Z}^k / \text{Im}(\varphi)|$  is finite the aforementioned property yields that

$$|\mathbb{Z}^k / \text{Im}(\varphi)| = \prod_{i=1}^k d_i.$$

Then by (\*) it follows that

$$|\mathbb{Z}^k / \text{Im}(\varphi)| = |\det(A)|.$$

This completes the proof.

## Result

3 of 3

Considering the Image of the map  $\varphi$  and using the diagonalisable property we have shown that  $|\mathbb{Z}^k / \text{Im}(\varphi)| = |\det(A)|$ .

7. a

To prove that there exist a matrix  $P \in GL_n(\mathbb{Z})$  such that  $PA = (d, 0, \dots, 0)^t$ ,

Suppose  $A = (a_1, a_2, \dots, a_n)^t$  be an integer column vector and  $d$  be the greatest common divisor of  $a_1, a_2, \dots, a_n$ .

Then, by Bezout's theorem, there exist some integers  $x_1, x_2, \dots, x_n$  such that,

$$x_1 a_1 + x_2 a_2 + \dots + x_n a_n = d$$

And there also exist some integers  $y_i, z_i, \dots, w_i \in \mathbb{Z}$  such that,

$$y_1 a_1 + y_2 a_2 + \dots + y_n a_n = 0, z_1 a_1 + z_2 a_2 + \dots + z_n a_n = 0, \dots, w_1 a_1 + w_2 a_2 + \dots + w_n a_n = 0$$

Where  $i = 1, 2, \dots, n$

Suppose  $GL_n(\mathbb{Z})$  be the general linear group of all  $n \times n$  invertible matrices.

Then, for any  $P \in GL_n(\mathbb{Z})$  such that,

$$P = \begin{bmatrix} x_1 & x_2 & \dots & \dots & x_n \\ y_1 & y_2 & \dots & \dots & y_n \\ z_1 & z_2 & \dots & \dots & z_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_1 & w_2 & \dots & \dots & w_n \end{bmatrix}, |P| \neq 0$$

Where,  $x_i, y_i, z_i, \dots, w_i \in \mathbb{Z}$

Then,

$$\begin{aligned} PA &= \begin{bmatrix} x_1 & x_2 & \dots & \dots & x_n \\ y_1 & y_2 & \dots & \dots & y_n \\ z_1 & z_2 & \dots & \dots & z_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_1 & w_2 & \dots & \dots & w_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 a_1 + x_2 a_2 + \dots + x_n a_n \\ y_1 a_1 + y_2 a_2 + \dots + y_n a_n \\ z_1 a_1 + z_2 a_2 + \dots + z_n a_n \\ \vdots \\ w_1 a_1 + w_2 a_2 + \dots + w_n a_n \end{bmatrix} \end{aligned}$$

Since,  $d$  be the greatest common divisor of  $a_1, a_2, \dots, a_n$ , then

$$x_1 a_1 + x_2 a_2 + \dots + x_n a_n = d$$

And

$$y_1 a_1 + y_2 a_2 + \dots + y_n a_n = 0, z_1 a_1 + z_2 a_2 + \dots + z_n a_n = 0, \dots, w_1 a_1 + w_2 a_2 + \dots + w_n a_n = 0$$

Now, substitute the above value in the matrix  $PA$ .

Then,

$$\begin{aligned} PA &= \begin{bmatrix} x_1 & x_2 & \dots & \dots & x_n \\ y_1 & y_2 & \dots & \dots & y_n \\ z_1 & z_2 & \dots & \dots & z_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_1 & w_2 & \dots & \dots & w_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 a_1 + x_2 a_2 + \dots + x_n a_n \\ y_1 a_1 + y_2 a_2 + \dots + y_n a_n \\ z_1 a_1 + z_2 a_2 + \dots + z_n a_n \\ \vdots \\ w_1 a_1 + w_2 a_2 + \dots + w_n a_n \end{bmatrix} \\ &= \begin{bmatrix} d \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= (d, 0, \dots, 0)^t \end{aligned}$$

Hence, it is proved that there exist a matrix  $P \in GL_n(\mathbb{Z})$  such that  $PA = (d, 0, \dots, 0)^t$ .

8. a

Consider the provided statement to diagonalize the matrix  $\begin{bmatrix} 3 & 2+i \\ 2-i & 9 \end{bmatrix}$ .

There are several steps to diagonalize the matrix is shown as below,

**Step1:** Interchange the value of column1 to column2 is as below;

$$\begin{bmatrix} 3 & 2+i \\ 2-i & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 2+i & 3 \\ 9 & 2-i \end{bmatrix}$$

**Step2:** column2 is modified as  $C_2 \rightarrow C_2 - C_1$  shown below;

$$\begin{bmatrix} 2+i & 3 \\ 9 & 2-i \end{bmatrix} \rightarrow \begin{bmatrix} 2+i & 3-(2+i) \\ 9 & 2-i-9 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 2+i & 1-i \\ 9 & -7-i \end{bmatrix}$$

**Step3:** Again interchange the column value that is  $C_1 \leftrightarrow C_2$

$$\begin{bmatrix} 2+i & 1-i \\ 9 & -7-i \end{bmatrix} \rightarrow \begin{bmatrix} 1-i & 2+i \\ -7-i & 9 \end{bmatrix}$$

**Step4:** In place of column2, first of all multiply column 1 with  $i$  and then subtract from column 2 is as below:

$$\begin{bmatrix} 1-i & 2+i \\ -7-i & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 1-i & 2+i+(1-i)i \\ -7-i & 9-i(-7-i) \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1-i & 2+i-(i+1) \\ -7-i & 9-(7i+1) \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1-i & 1 \\ -7-i & 8+7i \end{bmatrix}$$

**Step5:** Again replace both columns that is  $C_1 \leftrightarrow C_2$  shown as below;

$$\begin{bmatrix} 1-i & 1 \\ -7-i & 8+7i \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1-i \\ 8+7i & -7-i \end{bmatrix}$$

**Step6:** Perform operation  $C_2 \rightarrow (1-i)C_1 + C_2$  is shown as below;

$$\begin{bmatrix} 1 & 1-i \\ 8+7i & -7-i \end{bmatrix} \rightarrow \begin{bmatrix} 1 & (i-1)1+(1-i) \\ 8+7i & (i-1)(8+7i)+(-7-i) \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & i-1+1-i(i-1)1+(1-i) \\ 8+7i & (8i-7-8-7i)+(-7-i) \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 \\ 8+7i & -22 \end{bmatrix}$$

**Step7:** Perform  $R_2 \rightarrow (-8-7i)R_1 + R_2$  is shown as below;

$$\begin{bmatrix} 1 & 0 \\ 8+7i & -22 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ (-8-7i) \cdot 1 + (8+7i) & (-8-7i) \cdot 0 + (-22) \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 \\ (-8-7i+8+7i) & 0-22 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -22 \end{bmatrix}$$

Hence, the diagonalize matrix of the matrix  $\begin{bmatrix} 3 & 2+i \\ 2-i & 9 \end{bmatrix}$  is  $\begin{bmatrix} 1 & 0 \\ 0 & -22 \end{bmatrix}$ .

9. a

Consider the provided statement to prove that if  $L \subset M$  are lattices then  $[M : L] = \frac{\Delta(L)}{\Delta(M)}$

It is assumed that  $F_1, F_2$  are finite subsets of  $M$  then  $F_1 \subset F_2$ . Let  $x \in M$  be an arbitrary element. Then,

$$\begin{aligned}\Delta(M) &= \Delta(F) \\ &= \Delta(F \cup \{b\})\end{aligned}$$

[Comment](#)

Step 2 of 2 ^

As  $b \leftrightarrow \Delta(F \cup \{b\})$ ,  $\forall F \subset M$ , then from Lemma 8 as it is known that  $b \leftrightarrow \Delta(M)$

Hence  $M \leftrightarrow \Delta(M)$ , let  $b_1, b_2 \in M$ .

It is easy to show that,

$$\begin{aligned}\Delta(b_1 \wedge \Delta(M), b_2 \wedge \Delta(M)) &= \Delta(b_1, b_2) \wedge \Delta(M) \\ &= \Delta(M)\end{aligned}$$

From result, it is obtained that  $b_1 \wedge \Delta(M)$  and  $b_2 \wedge \Delta(M)$  is compatible in  $[0, \Delta(M)]$ .

Because, this is lattice logic which is regular then  $[M : L] = \frac{\Delta(L)}{\Delta(M)}$  is **proved**.

## Section 5

1. a

Consider the provided statement to determine a presentation matrix as  $R$ -module for the ideal  $(2, 1+\delta)$  where  $\delta = \sqrt{-5}$ .

[Comment](#)

Step 2 of 2 ^

The surjective map is considered as shown below;

$$\begin{aligned}R^2 &\rightarrow \phi(2, 1+\delta) \\ (x, y) &\rightarrow 2x + (1+\delta)y\end{aligned}$$

As  $\ker \phi$  has two generators. Then, there is the relation that  $2(-1-\delta) + (1+\delta)2 = 0$  and also the relation  $2(-3) + (1+\delta)(1-\delta) = 0$ .

Therefore, two relations cannot be derived from each other. Then  $\ker \phi$  is find by finding  $(x, y) \in \mathbb{R}$  such that  $2x + (1+\delta)y = 0$ .

It means that  $\ker \phi = \left\{ y : y \in \mathbb{R} \text{ and } \left(\frac{1+\delta}{2}\right)y \text{ is also in } \mathbb{R} \right\}$

Therefore, the presentation matrix is  $\begin{bmatrix} -3 & -1-\delta \\ 1-\delta & 2 \end{bmatrix}$ .

2. a



Consider the provided statement to identify the Abelian group that is presented by the matrix is provided as below;

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 3 & 6 \end{bmatrix}$$

For determining the Abelian group corresponding to a presentation matrix, the provided matrix is changed into Smith normal form is as below;

**Step 1:** Interchange  $R_1 \leftrightarrow R_2$  is shown as below;

$$\begin{bmatrix} 3 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 3 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 2 \\ 2 & 3 & 6 \end{bmatrix}$$

**Step2:** Interchange  $R_2 \leftrightarrow R_3$  is shown as below;

$$\begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 2 \\ 2 & 3 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 6 \\ 3 & 1 & 2 \end{bmatrix}$$

**Step3:** Perform  $R_2 \rightarrow R_2 - 2 \cdot R_1$  is shown as below;

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 6 \\ 3 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 2-2 & 3-2 & 6-2 \\ 3 & 1 & 2 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 3 & 1 & 2 \end{bmatrix}$$

**Step4:** Perform  $R_3 \rightarrow R_3 - 3 \cdot R_1$  is shown as below;

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 3 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 3-3 & 1-3 & 2-3 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix}$$

**Step5:** Perform  $R_1 \rightarrow R_1 + R_3$  is shown as below;

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1-2 & 1-1 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix}$$

**Step6:** Perform  $C_2 \rightarrow C_1 + C_2$  is shown as below;

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1-1 & 0 \\ 0 & 1+0 & 4 \\ 0 & -2+0 & -1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix}$$

**Step7:** Perform  $C_3 \rightarrow C_3 - 4C_2$  is shown as below;

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4-4 \\ 0 & -2 & -1+8 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 7 \end{bmatrix}$$

**Step8:** Perform  $R_3 \rightarrow R_3 + 2R_2$  is shown as below;

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2+2 & 7+0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Hence, the Abelian group of the presented matrix is of group  $G \cong \boxed{\frac{\mathbb{Z}}{7}}$ .

## Section 6

1. a

Consider the provided statement to prove that there is a finite subset of these polynomials whose zeros define the same locus. As it is provided that let  $V \subseteq \mathbb{C}^n$  is the locus of common zeros for the infinite set of polynomials  $f_1, f_2, f_3, \dots$ .

[Comment](#)

**Step 2 of 2** ^

Let it is known that  $V$  is a variety and it can be associate that an ideal of polynomials to this variety.

Let  $I = (f_1, f_2, \dots)$  then the variety associated with  $I, V(I)$  is precisely  $V(I) = V$ .

$$(f_1) \subseteq (f_1, f_2) \subseteq (f_1, f_2, f_3) \subseteq \dots$$

$$V(f_1) \supseteq V(f_1, f_2) \supseteq V(f_1, f_2, f_3) \supseteq \dots$$

As  $\mathbb{C}^n$  is Noetherian and there can be no infinitely ascending chains of ideals and there exists  $n$  such that  $(f_1, \dots, f_n) = I$  and  $V(f_1, \dots, f_n) = V$ .

Hence, provided statement is **proved**.

2. a

Consider the provided statement to find a ring  $R$  that is not finitely generated.

An example of the ring is  $R = F[x_1, x_2, x_3, \dots]$  a polynomial ring in a countably infinite number of variables. From the Hilbert Basis Theorem,

A polynomial ring with a finite number of variables is Noetherian; therefore all ideals would be finitely generated.

[Comment](#)

### Step 2 of 3 ^

Let  $I$  is the ideal generated by all the variables. It can be check that  $I$  is an ideal by showing that for  $a, b \in I, r \in R$  then  $a + b$  and  $ra \in I$ . Where  $x_i \in I, \forall i \in \mathbb{N}$  and  $1 \notin I$ , therefore the smallest basis for  $I$  is  $(x_1, x_2, \dots)$  an infinite collection of all the variables.

The span of any smaller basis would not include the element  $x_i$  for some  $x_i$  since there are no relations that allow  $x_i$  to be written as a  $R$ -linear combination of other variables.

There is also an example  $R = \{\text{Continuous function from } \mathbb{R} \rightarrow \mathbb{R}\}$ . This is a ring that using  $(f + g)(x) = f(x) + g(x)$  and  $(f \cdot g)(x) = f(x) \cdot g(x)$ . Remember that addition and multiplication preserve continuity that is,

$$I = \{f : f(0) = 0\}$$

To check that  $I$  is an ideal it is noted that if  $f(0) = 0$  and  $g(0) = 0$  then

$$\begin{aligned}(f + g)(0) &= f(0) + g(0) \\ &= 0\end{aligned}$$

And also,

$$\begin{aligned}(f \cdot g)(0) &= f(0) \cdot g(0) \\ &= 0\end{aligned}$$

If it is to be seen that,  $I$  is not finitely generated as  $\sin x, \sin 2x, \sin 3x, \dots$  all are linearly independent and also have no common divisor but are all in  $I$ . Hence, these all are examples of a ring that is **not finitely generated**.

## Section 7

1. a

Consider the provided statement to find the direct sum of cyclic groups isomorphic to the Abelian group. Provided matrix is,

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

Now diagonalize the provided matrix in various steps which is shown as below,

**Step1:** Perform  $C_2 \rightarrow C_2 - C_1$  and  $C_3 \rightarrow C_3 - C_1$  is as below;

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 2-2 & 2-2 \\ 2 & 2-2 & 0-2 \\ 2 & 0-2 & 2-2 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 2 & 0 & -2 \\ 2 & -2 & 0 \end{bmatrix}$$

**Step2:** Perform  $C_1 \rightarrow C_1 + C_2$

$$\begin{bmatrix} 2 & 0 & 0 \\ 2 & 0 & -2 \\ 2 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2+0 & 0 & 0 \\ 2-2 & 0 & -2 \\ 2+0 & -2 & 0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -2 \\ 2 & -2 & 0 \end{bmatrix}$$

**Step3:** Perform  $C_1 \rightarrow C_1 + C_2$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -2 \\ 2 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2+0 & 0 & 0 \\ 0+0 & 0 & -2 \\ 2-2 & -2 & 0 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & -2 & 0 \end{bmatrix}$$

**Step4:** Perform  $C_2 \leftrightarrow C_3$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Therefore, this is diagonalise form of the provided matrix. Therefore, the Abelian group presented by the matrix is the direct sum of three cyclic groups that is  $C_2 \oplus C_2 \oplus C_2$ .

2. a

Consider the provided statement to write the Abelian group which is generated by  $x$  and  $y$  with the given relation.

[Comment](#)

Step 2 of 3 ^

Provided relation is,

$$3x + 4y = 0$$

This equation can be also written in the form of,  $AX = 0$  where  $A, X$  are matrices.

$$\begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Where  $A = \begin{bmatrix} 3 & 4 \end{bmatrix}$  and  $X = \begin{bmatrix} x \\ y \end{bmatrix}$

The presented matrix  $A = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$  can be diagonalized as shown below;

**Step1:**  $R_2 \rightarrow 4 \cdot R_2 - 5 \cdot R_1$  is as below;

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} \rightarrow \begin{bmatrix} 3 \\ 4 \cdot 4 - 5 \cdot 3 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 3 \\ 16 - 15 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

**Step2:**  $R_1 \rightarrow R_1 - 3 \cdot R_2$  is as below;

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 3 - 3 \cdot 1 \\ 1 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 3 - 3 \\ 1 \end{bmatrix} \\ \rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The obtained matrix which is in diagonalized form is  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ . Hence, this represents the free Abelian group whose rank is 1 because as from the definition of rank, the number of non-zero row is 1, so the rank is 1.

### 3. a

(a)

To find an isomorphic direct product of cyclic groups for the abelian group  $V$  generated by  $x, y, z$ ,

Suppose  $V$  be the abelian group generated by  $x, y, z$  with the following relation,

$$3x + 2y + 8z = 0, 2x + 4z = 0$$

And a trivial relation

$$0x + 0y + 0z = 0$$

Then, the system of relation is,

$$3x + 2y + 8z = 0, 2x + 4z = 0, 0x + 0y + 0z = 0$$

First, find the diagonal matrix corresponding to the system of relations of the abelian group  $V$ .

Since, the matrix corresponding to the system of relations is,

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 0 & 0 \\ 8 & 4 & 0 \end{bmatrix}$$

Suppose  $\lambda$  be the eigenvalue of the matrix  $A$ , then

$$|A - \lambda I| = 0 \\ \begin{vmatrix} 3 - \lambda & 2 & 0 \\ 2 & 0 - \lambda & 0 \\ 8 & 4 & 0 - \lambda \end{vmatrix} = 0$$

$$(-\lambda)[(3 - \lambda)(-\lambda) - 4] = 0$$

$$(-\lambda)[\lambda^2 - 3\lambda - 4] = 0$$

$$\lambda = 0, -1, 4$$

Therefore, the eigenvalues of the matrix  $A$  are  $0, -1, 4$

Now, find the eigenvector corresponding to each eigenvalue of the matrix  $A$ .

Suppose  $v$  be the eigenvector corresponding to the eigenvalue  $0$  such that

$$v = (x_1, x_2, x_3)^T$$

Then, for the eigenvector of the matrix  $A$

$$[A - \lambda I](v) = 0$$

$$[A - 0I](v) = 0$$

$$\begin{bmatrix} 3-0 & 2 & 0 \\ 2 & 0-0 & 0 \\ 8 & 4 & 0-0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

Then, the systems of equations are,

$$3x_1 + 2x_2 = 0, 2x_1 = 0, 8x_1 + 4x_2 = 0$$

Then,

$$x_1 = 0, x_2 = 0, x_3 = 1$$

Then, the eigenvector corresponding to the eigenvalue  $0$  is,

$$v = (0, 0, 1)^T$$

In the similar way, the eigenvector corresponding to the eigenvalue  $-1$  is  $\left(-\frac{1}{2}, 1, 0\right)$  and the

eigenvector corresponding to the eigenvalue  $4$  is  $\left(\frac{2}{5}, \frac{1}{5}, 1\right)$

Then, the matrix  $P$  of the eigenvectors corresponding to the eigenvalues is,

$$P = \begin{bmatrix} 0 & -\frac{1}{2} & \frac{2}{5} \\ 0 & 1 & \frac{1}{5} \\ 1 & 0 & 1 \end{bmatrix}$$

Then, the diagonal matrix corresponding to the matrix  $A$  is,

$$D = P^{-1}AP$$

$$\begin{aligned} &= \begin{bmatrix} -2 & -1 & 1 \\ \frac{2}{5} & \frac{4}{5} & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 2 & 0 & 0 \\ 8 & 4 & 0 \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{2} & \frac{2}{5} \\ 0 & 1 & \frac{1}{5} \\ 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$

Therefore, the diagonal matrix associated with the matrix  $A$  is,

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Then, by the structural theorem for the abelian group  $V$  is,

$$V \cong \{0\} \oplus C_1 \oplus C_4$$

Where,  $\{0\}$  is free abelian group.

Hence, the required isomorphic direct product of cyclic groups for the abelian group  $V$  is

$$\boxed{\{0\} \oplus C_1 \oplus C_4}.$$



(b)

To find an isomorphic direct product of cyclic groups for the abelian group  $V$  generated by  $x, y, z$ ,

Suppose  $V$  be the abelian group generated by  $x, y, z$  with the following relation,

$$x + y = 0, 2x = 0, 4x + 2z = 0, 4x + 2y + 2z = 0$$

Then, the system of linearly independent relation is,

$$x + y = 0, 4x + 2z = 0, 4x + 2y + 2z = 0$$

First, find the diagonal matrix corresponding to the system of relations of the abelian group  $V$ .

Since, the matrix corresponding to the system of relations is,

$$B = \begin{bmatrix} 1 & 4 & 4 \\ 1 & 0 & 2 \\ 0 & 2 & 2 \end{bmatrix}$$

Suppose  $\lambda$  be the eigenvalue of the matrix  $B$ , then

$$\begin{aligned} |B - \lambda I| &= 0 \\ \begin{vmatrix} 1-\lambda & 4 & 4 \\ 1 & 0-\lambda & 2 \\ 0 & 2 & 2-\lambda \end{vmatrix} &= 0 \\ -\lambda^3 + 3\lambda^2 + 6\lambda - 4 &= 0 \\ \lambda &= 2, 1 + \sqrt{5}, 1 - \sqrt{5} \end{aligned}$$

Then, the eigenvalues of the matrix  $B$  are  $2, 1 + \sqrt{5}, 1 - \sqrt{5}$ .

Therefore, the diagonal matrix associated with the matrix  $B$  is,

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 + \sqrt{5} & 0 \\ 0 & 0 & 1 - \sqrt{5} \end{bmatrix}$$

Then, by the structural theorem for the abelian group  $V$  is,

$$V \cong C_2 \oplus \mathbb{R} \oplus \mathbb{R}$$

Where,  $\mathbb{R}$  is free abelian group.

Hence, the required isomorphic direct product of cyclic groups for the abelian group  $V$  is

$$\boxed{C_2 \oplus \mathbb{R} \oplus \mathbb{R}}.$$

(c)

To find an isomorphic direct product of cyclic groups for the abelian group  $V$  generated by  $x, y, z$ ,

Suppose  $V$  be the abelian group generated by  $x, y, z$  with the following relation,

$$2x + y = 0, x - y + 3z = 0$$

And a trivial relation

$$0x + 0y + 0z = 0$$

Then, the system of relation is,

$$2x + y = 0, x - y + 3z = 0, 0x + 0y + 0z = 0$$

First, find the diagonal matrix corresponding to the system of relations of the abelian group  $V$ .

Since, the matrix corresponding to the system of relations is,

$$C = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

Suppose  $\lambda$  be the eigenvalue of the matrix  $C$ , then

$$\begin{aligned} |C - \lambda I| &= 0 \\ \begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & -1-\lambda & 0 \\ 0 & 3 & 0-\lambda \end{vmatrix} &= 0 \\ (-\lambda)[(2-\lambda)(-1-\lambda)-1] &= 0 \\ (-\lambda)(\lambda^2 - \lambda - 3) &= 0 \\ \lambda &= 0, \frac{1}{2}(1-\sqrt{13}), \frac{1}{2}(1+\sqrt{13}) \end{aligned}$$

Then, the eigenvalues of the matrix  $C$  are  $0, \frac{1}{2}(1-\sqrt{13}), \frac{1}{2}(1+\sqrt{13})$ .

Therefore, the diagonal matrix associated with the matrix  $C$  is,

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2}(1-\sqrt{13}) & 0 \\ 0 & 0 & \frac{1}{2}(1+\sqrt{13}) \end{bmatrix}$$

Then, by the structural theorem for the abelian group  $V$  is,

$$V \cong \{0\} \oplus \mathbb{R} \oplus \mathbb{R}$$

Where,  $\{0\}$  and  $\mathbb{R}$  is free abelian group.

Hence, the required isomorphic direct product of cyclic groups for the abelian group  $V$  is

$$\boxed{\{0\} \oplus \mathbb{R} \oplus \mathbb{R}}.$$

(d)

To find an isomorphic direct product of cyclic groups for the abelian group  $V$  generated by  $x, y, z$ ,

Suppose  $V$  be the abelian group generated by  $x, y, z$  with the following relation,

$$7x + 5y + 2z = 0, 3x + 3y = 0, 13x + 11y + 2z = 0$$

First, find the diagonal matrix corresponding to the system of relations of the abelian group  $V$ .

Since, the matrix corresponding to the system of relations is,

$$D = \begin{bmatrix} 7 & 3 & 13 \\ 5 & 3 & 11 \\ 2 & 0 & 2 \end{bmatrix}$$

Suppose  $\lambda$  be the eigenvalue of the matrix  $D$ , then

$$\begin{aligned} |D - \lambda I| &= 0 \\ \begin{vmatrix} 7-\lambda & 3 & 13 \\ 5 & 3-\lambda & 11 \\ 2 & 0 & 2-\lambda \end{vmatrix} &= 0 \\ 12\lambda^2 - \lambda^3 &= 0 \\ \lambda &= 0, 0, 12 \end{aligned}$$

Then, the eigenvalues of the matrix  $D$  are  $0, 0, 12$ .

Since,

$$G.M(\lambda) = n - \text{rank}(D - \lambda I)$$

Then,

$$\begin{aligned} G.M(0) &= 3 - \text{rank}(D - 0I) \\ &= 3 - 2 \\ &= 1 \end{aligned}$$

Then, the geometric multiplicity of eigenvalue 0 is 1.

But, the algebraic multiplicity of eigenvalue 0 is 2.

This implies that, the matrix  $D$  is not diagonalizable.

Then, by the structural theorem for the infinitely generated abelian group  $V$  is,

$$V \cong \mathbb{R} \oplus \mathbb{R} \oplus \mathbb{R}$$

Where,  $\mathbb{R}$  is free abelian group.

Hence, the required isomorphic direct product of cyclic groups for the abelian group  $V$  is

$$\mathbb{R} \oplus \mathbb{R} \oplus \mathbb{R}.$$

4. a

To identify the abelian group corresponding to the presentation matrix,

Consider the following presentation matrix,

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Then, the presentation matrix can be diagonalized by the elementary row operations.

Then, apply the elementary row operation  $R_1 \leftarrow R_1 - 2R_2$  in the presentation matrix.

Then,

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This implies that, the presentation matrix corresponds to the free abelian group of the rank 1.

Then, the abelian group corresponding to the presentation matrix is,

$$\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \mathbb{Z}$$

Hence, the required abelian group corresponding to the presentation matrix is free abelian group

$$\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \mathbb{Z}.$$

Consider the following presentation matrix,

$$\begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

This implies that, the presentation matrix corresponds to the free abelian group of the rank 1.

Suppose  $(u_1, u_2)$  be the generator of the group corresponding to the presentation matrix.

Then, the relation for group is,

$$5u_2 = 0$$

Then, the group corresponding to the presentation matrix is,

$$\mathbb{Z} \oplus \frac{\mathbb{Z}}{5\mathbb{Z}}$$

Then, the abelian group corresponding to the presentation matrix is,

$$\mathbb{Z} \oplus \frac{\mathbb{Z}}{5\mathbb{Z}}$$

Hence, the required abelian group corresponding to the presentation matrix is abelian group

$$\mathbb{Z} \oplus \frac{\mathbb{Z}}{5\mathbb{Z}}.$$

Consider the following presentation matrix,

$$\begin{bmatrix} 2 & 0 & 0 \end{bmatrix}$$

This implies that, the presentation matrix corresponds to the free abelian group of the rank 1.

Suppose  $(u_1, u_2, u_3)$  be the generator of the group corresponding to the presentation matrix.

Then, the relation for group is,

$$2u_1 = 0$$

Then, the group corresponding to the presentation matrix is,

$$\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \mathbb{Z} \oplus \mathbb{Z}$$

Hence, the required abelian group corresponding to the presentation matrix is abelian group

$$\boxed{\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \mathbb{Z} \oplus \mathbb{Z}}.$$

Consider the following presentation matrix,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

This implies that, the presentation matrix corresponds to the free abelian group of the rank 2.

Suppose  $(u_1, u_2, u_3)$  be the generator of the group corresponding to the presentation matrix.

Then, the first relation for group is,

$$u_1 = 0$$

And the second relation for the group is,

$$u_2 = 0$$

Therefore, eliminate the zero elements from the generating set  $(u_1, u_2, u_3)$ .

Then, in the generating set, there is only one element  $\{u_3\}$ .

Then, the abelian group corresponding to the presentation matrix is,

$$\mathbb{Z}$$

Hence, the required abelian group corresponding to the presentation matrix is free abelian group

$$\boxed{\mathbb{Z}}.$$

Consider the following presentation matrix,

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

Then, the presentation matrix corresponds to the free abelian group of the rank 2.

First, find the diagonal matrix corresponding the above presentation matrix.

Suppose  $\lambda$  be the eigenvalue of the matrix  $A$ , then

$$\begin{aligned} |A - \lambda I| &= 0 \\ \begin{vmatrix} 2-\lambda & 3 \\ 1 & 2-\lambda \end{vmatrix} &= 0 \\ (2-\lambda)^2 - 3 &= 0 \\ \lambda^2 - 4\lambda + 4 - 3 &= 0 \\ \lambda^2 - 4\lambda + 1 &= 0 \\ \lambda &= 2 + \sqrt{3}, 2 - \sqrt{3} \end{aligned}$$

Then, the eigenvalues of the matrix  $A$  are  $2 + \sqrt{3}, 2 - \sqrt{3}$ .

Therefore, the diagonal matrix associated with the matrix  $A$  is,

$$D = \begin{bmatrix} 2 + \sqrt{3} & 0 \\ 0 & 2 - \sqrt{3} \end{bmatrix}$$

Then, by the structural theorem for the abelian group corresponding to the presentation matrix is,

$$\mathbb{R} \oplus \mathbb{R}$$

Where,  $\mathbb{R}$  is free abelian group.

Hence, the required abelian group corresponding to the presentation matrix is free abelian group

$$\boxed{\mathbb{R} \oplus \mathbb{R}}.$$

Consider the following presentation matrix,

$$B = \begin{bmatrix} 2 & 4 \\ 1 & 4 \end{bmatrix}$$

Then, the presentation matrix corresponds to the free abelian group of the rank 2.

First, find the diagonal matrix corresponding the above presentation matrix.

Suppose  $\lambda$  be the eigenvalue of the matrix  $B$ , then

$$\begin{aligned} |B - \lambda I| &= 0 \\ \begin{vmatrix} 2-\lambda & 4 \\ 1 & 4-\lambda \end{vmatrix} &= 0 \\ \lambda^2 - 6\lambda + 8 - 4 &= 0 \\ \lambda^2 - 6\lambda + 4 &= 0 \\ \lambda &= 3 + \sqrt{5}, 3 - \sqrt{5} \end{aligned}$$

Then, the eigenvalues of the matrix  $B$  are  $3 + \sqrt{5}, 3 - \sqrt{5}$ .

Therefore, the diagonal matrix associated with the matrix  $B$  is,

$$D = \begin{bmatrix} 3 + \sqrt{5} & 0 \\ 0 & 3 - \sqrt{5} \end{bmatrix}$$

Then, by the structural theorem for the abelian group corresponding to the presentation matrix is,

$$\mathbb{R} \oplus \mathbb{R}$$

Where,  $\mathbb{R}$  is free abelian group.

Hence, the required abelian group corresponding to the presentation matrix is free abelian group

$$\boxed{\mathbb{R} \oplus \mathbb{R}}.$$

Consider the following presentation matrix,

$$C = \begin{bmatrix} 2 & 4 \\ 6 & 4 \end{bmatrix}$$

Then, the presentation matrix corresponds to the free abelian group of the rank 2 .

First, find the diagonal matrix corresponding the above presentation matrix.

Suppose  $\lambda$  be the eigenvalue of the matrix  $C$  , then

$$\begin{aligned} |C - \lambda I| &= 0 \\ \begin{vmatrix} 2-\lambda & 4 \\ 6 & 4-\lambda \end{vmatrix} &= 0 \\ \lambda^2 - 6\lambda + 8 - 24 &= 0 \\ \lambda^2 - 6\lambda - 16 &= 0 \\ \lambda^2 - 8\lambda + 2\lambda - 16 &= 0 \\ (\lambda - 8)(\lambda + 2) &= 0 \\ \lambda &= -2, 8 \end{aligned}$$

Then, the eigenvalues of the matrix  $C$  are  $-2, 8$  .

Therefore, the diagonal matrix associated with the matrix  $C$  is,

$$D = \begin{bmatrix} -2 & 0 \\ 0 & 8 \end{bmatrix}$$

Then, by the structural theorem for the abelian group, the abelian group corresponding to the presentation matrix is,

$$C_2 \oplus C_8$$

Hence, the required abelian group corresponding to the presentation matrix is free abelian group

$$\boxed{C_2 \oplus C_8}.$$


---

Consider the following presentation matrix,

$$E = \begin{bmatrix} 4 & 6 \\ 2 & 3 \end{bmatrix}$$

Then, the presentation matrix corresponds to the free abelian group of the rank 2 .

First, find the diagonal matrix corresponding the above presentation matrix.

Suppose  $\lambda$  be the eigenvalue of the matrix  $E$  , then

$$\begin{aligned} |E - \lambda I| &= 0 \\ \begin{vmatrix} 4-\lambda & 6 \\ 2 & 3-\lambda \end{vmatrix} &= 0 \\ \lambda^2 - 7\lambda + 12 - 12 &= 0 \\ \lambda^2 - 7\lambda &= 0 \\ \lambda &= 0, 7 \end{aligned}$$

Then, the eigenvalues of the matrix  $E$  are  $0, 7$  .

Therefore, the diagonal matrix associated with the matrix  $E$  is,

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 7 \end{bmatrix}$$

Then, by the structural theorem for the abelian group, the abelian group corresponding to the presentation matrix is,

$$\mathbb{Z} \oplus \frac{\mathbb{Z}}{7\mathbb{Z}}$$

Where,  $\mathbb{Z}$  is free abelian group.

Hence, the required abelian group corresponding to the presentation matrix is free abelian group

$$\boxed{\mathbb{Z} \oplus \frac{\mathbb{Z}}{7\mathbb{Z}}}.$$

5. a



Consider the provided statement to determine the number of isomorphism classes of abelian groups of order 400.

Comment

Step 2 of 3 ^

If the group is isomorphic, then to a sum of  $C_{d_1} \oplus C_{d_2} \oplus \dots \oplus C_{d_k}$  with  $d_1 | d_2 | \dots | d_k$  and  $d_1 d_2 \dots d_k = 400$ . As it is known that the prime factorization of 400 is  $400 = 2^4 5^2$ . Therefore, the values of  $k$  is less than or equal to 4.

The choices for  $d_1$  are divisors of  $400 \leq 20$  or just 400. Moreover, it can't have  $2^3 | d_1$  or  $5^2 | d_1$ , so the value of  $d_1$  is 2, 4, 5, 10, 20. If  $5 | d_1$  then there is  $d_3$  because there are only two copies of 5 to go ground.

Therefore,

$$d_2 = \frac{400}{d_1}.$$

That gives three equivalence types  $C_5 \oplus C_{80}, C_{10} \oplus C_{40}, C_{20} \oplus C_{20}$

If  $d_1 = 2$ , it could have  $d_2 = 200$  to get  $C_2 \oplus C_{200}$ , otherwise  $4 | d_2$  because  $2 \times 4 \nmid 400$ . Therefore, there are two options  $d_2 = 2$  or  $d_2 = 10$ . If  $d_2 = 10$  then the factors of 5 constrain  $d_3 = 20$  and the group is  $C_2 \oplus C_{10} \oplus C_{20}$ .

If  $d_1 = d_2 = 2$  then  $d_3 = 100$  and to find  $C_2 \oplus C_2 \oplus C_{100}$  or as  $4 | d_3$  that is  $d_3 = 10$  and  $d_4 = 10$  to find  $C_2 \oplus C_2 \oplus C_{10} \oplus C_{10}$  or  $d_3 = 2$  and  $d_4 = 50$  to find  $C_2 \oplus C_2 \oplus C_2 \oplus C_{50}$ .

Hence, the number of isomorphism is 10.

6. a

(a)

Consider the provided statement to prove that cyclic group  $C_{ab}$  is isomorphic to the product  $C_a \oplus C_b$ .

Let  $a$  and  $b$  are relatively prime integers then a cyclic group of order  $ab$  is isomorphic to the product of a cyclic group of order  $a$  and a cyclic group of order  $b$ .

On other side, a cyclic group of order 4 is not isomorphic to a product of two cyclic groups of order 2. Every element of  $C_2 \times C_2$  has order 1 or 2 but a cyclic group of order 4 contains two elements of order 4.

Therefore, if  $a$  and  $b$  are relatively prime integers then the cyclic group  $C_{ab}$  that has order  $ab$  is isomorphic to the direct sum  $C_a \oplus C_b$  of cyclic subgroups of orders  $a$  and  $b$  is **proved**.

(b)

Consider the provided statement to explain what happen if  $a$  and  $b$  are relatively prime is dropped.

When  $a$  and  $b$  are not relatively prime integers then their greatest common divisor is not 1 then it will also not follow the property as  $ar + bs \neq 1$ .

Let  $a$  and  $b$  are not relatively prime integers then a cyclic group of order  $ab$  is not isomorphic to the product of a cyclic group of order  $a$  and a cyclic group of order  $b$  is not the direct sum  $C_a \oplus C_b$  of cyclic subgroups of orders  $a$  and  $b$ .

7. a

Consider the provided statement to write the provided module as a direct sum of cyclic modules. The elements of  $V$  are  $v_1$  and  $v_2$ , relations are as  $(1+i)v_1 + (2-i)v_2 = 0$  and  $3v_1 + 5iv_2 = 0$

To find the presentation matrix  $A$  use the provided relation and then diagonalize it.

$$A = \begin{bmatrix} 1+i & 3 \\ 2-i & 5i \end{bmatrix}$$

Then,

$$\begin{bmatrix} 1+i & 3 \\ 2-i & 5i \end{bmatrix} \xrightarrow{\begin{pmatrix} 1 & 0 \\ -1+i & 1 \end{pmatrix}} \begin{bmatrix} 1+i & 3 \\ -i & -3+8i \end{bmatrix}$$

Further use identity matrix is shown as below,

$$\begin{bmatrix} 1+i & 3 \\ -i & -3+8i \end{bmatrix} \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \begin{bmatrix} -i & -3+8i \\ 1+i & 3 \end{bmatrix}$$

Now when the matrix  $\begin{bmatrix} i & 0 \\ 1 & 0 \end{bmatrix}$  is used, then

$$\begin{bmatrix} -i & -3+8i \\ 1+i & 3 \end{bmatrix} \xrightarrow{\begin{pmatrix} i & 0 \\ 1 & 0 \end{pmatrix}} \begin{bmatrix} 1 & -8-3i \\ 1+i & 3 \end{bmatrix}$$

Now the matrix  $\begin{bmatrix} 1 & 0 \\ -1-i & 1 \end{bmatrix}$  is used, then

$$\begin{bmatrix} 1 & -8-3i \\ 1+i & 3 \end{bmatrix} \xrightarrow{\begin{pmatrix} 1 & 0 \\ -1-i & 1 \end{pmatrix}} \begin{bmatrix} 1 & -8-3i \\ 0 & 8+11i \end{bmatrix}$$

The obtained result reduces to the  $1 \times 1$  matrix  $(8+11i)$  by using proposition 14.5.7. Therefore,

$$V \approx \frac{R}{(8+11i)}$$

Now, to decompose further

$$\begin{aligned} N(8+11i) &= 64+121 \\ &= 185 \\ &= 5 \cdot 37 \\ &= (2+i)(2-i)(6+i)(6-i) \end{aligned}$$

It is also known that,

$$\begin{aligned} i(2-i)(6-i) &= i(12-8i-1) \\ &= 8+11i \end{aligned}$$

Therefore,

$$\begin{aligned} V &\approx \frac{R}{(8+11i)} \\ &\approx \frac{R}{(2-i)} \oplus \frac{R}{(6-i)} \end{aligned}$$

Hence  $V$  can be written as a sum of cyclic modules,  $V \approx \boxed{\frac{R}{(2-i)} \oplus \frac{R}{(6-i)}}$ .

8. a

If  $X$  is defined as a  $\mathbb{Z}_-$ -module, then this implies that  $(X, +)$  is an abelian group with the addition property.

Consider  $\mathbb{F} = \mathbb{F}_p$

Take any prime  $p \in \mathbb{Z}$

Now, consider the quotient ring  $\mathbb{Z}$  with the modulus ideal  $p\mathbb{Z}$

That is;

$$\mathbb{Z}/p\mathbb{Z}$$

Now, since;

$$\mathbb{F}_p = \{0, 1, \dots, p-1\}$$

The operations used on this group will be addition and multiplication of integer  $\text{mod } p$

Further, for every,  $x \in \mathbb{F}_p^*$  has a multiplicative inverse.

Since,  $x \in \mathbb{F}_p^*$  and here  $p$  is a prime

Then, with  $\mathbb{Z}[i]$ -module over the addition group  $F^+$ ;

$$\gcd(x, p) = 1$$

Hence, by using Euclid it can be obtained that there exists  $a, b$  such that;

$$xa + pb = 1$$

It will be equal to 1 as, the additive group is over  $F^+$ , and that is mod 1.

Next taking the additive group  $F^+$ , that is  $\text{mod } 2$ , the expression becomes;

$$xa + pb = 2$$

9. a

Consider the provided statement to prove that provided concepts are equivalent.

Provided condition is,  $R$ -module where  $R = \mathbb{Z}[i]$  and there is an abelian group  $V$  with a homomorphism  $\phi: V \rightarrow V$  such that  $\phi \circ \phi = -\text{identity}$ .

[Comment](#)

Step 2 of 3 ^

Let  $X$  be a  $\mathbb{Z}[i]$  module. Then  $X$  has the structure of an abelian group under addition and a function  $\phi$  such that  $\phi: X \rightarrow X, \phi(m) = im$  is a homomorphism that satisfying the condition  $\phi^2 = -\text{identity}$ . It is note that  $\phi': x \mapsto -ix$  is another homomorphism.

Conversely, it is assumed that  $V$  be an abelian group with such a homomorphism  $\phi$ . Then  $V$  has the structure of a  $\mathbb{Z}$  module such that,

$$nv = v + v + \dots + v$$

Where  $n$  are the summands, now the structure of a  $\mathbb{Z}[i]$  module is defined as below:

$$(a + bi)v = av + b\phi(v)$$

Now it is checked that this is appropriately distributive and associative:

$$\begin{aligned} (a + bi)(u + v) &= a(u + v) + b\phi(u + v) \\ &= au + av + b\phi(u) + b\phi(v) \\ &= (aub\phi(u)) + (av + b\phi(v)) \end{aligned}$$

As it is provided that  $\phi$  is a homomorphism; then

$$\begin{aligned} ((a+bi)+(c+di))v &= (a+c)v + (b+d)\phi(v) \\ &= av + cv + b\phi(v) + d\phi(v) \\ &= (av + b\phi(v)) + (cv + d\phi(v)) \end{aligned}$$

For the associativity of multiplication, only need to look at purely real or purely imaginary terms and having dealt with distributive. Anything involving purely real terms will work, because  $\phi$  is a homomorphism of abelian groups.

It is just needed to check that  $i(iv) = -v$  that corresponds precisely to  $\phi^2 = -\text{identity}$ .

It can be note that  $\phi$  can be mapped to multiplication instead of  $-i$ . Hence, provided statement is **proved**.

## Section 8

1. a

Consider the provided statement to check that  $\mathbb{C}[t]$ -module cyclic.

It is assumed that the provided matrix is  $A$  and the value is  $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$ . It is also providing that  $T$  is the linear operator on  $\mathbb{C}^2$ .

[Comment](#)

Step 2 of 3 ^

The characteristic polynomial of the provided matrix is given as,

$$f(t) = \det(A - tI)$$

Therefore,

$$\begin{aligned} f(t) &= \begin{vmatrix} 2-t & 1 \\ 0 & 1-t \end{vmatrix} = \begin{vmatrix} 2-t & 1 \\ 0 & 1-t \end{vmatrix} \\ &= \begin{vmatrix} 2-t & 1 \\ 0 & 1-t \end{vmatrix} \\ &= (2-t)(1-t) \\ &= -(t-2)(1-t) \end{aligned}$$

Since, the characteristic polynomial is  $(t-2)(t-1)$ , it is relatively prime. The module is also

isomorphic to  $\frac{\mathbb{C}[t]}{((t-2)(t-1))}$ .

As an element generates the module, if and only if it does not belong to a maximal non-trivial sub-module, as the provided matrix is 2-dimensional as a vector space. Therefore the non-trivial maximal sub-modules must be 1-dimensional so it is needed to be Eigen space of  $T$ . Hence it generates all of  $\mathbb{C}^2$ . Therefore, it is clearly  $\mathbb{C}[t]$ -module cyclic.

2. a

Consider the provided statement to show that the matrix  $M$  of the corresponding linear operator is a Jordan block. As it is provided that  $M$  is a  $\mathbb{C}[t]$  module of form  $\frac{\mathbb{C}[t]}{(t-\alpha)^n}$ .

[Comment](#)

### Step 2 of 3 ^

From the proposition 11.5.5, the basis for  $M$  over  $\mathbb{C}$  is  $1, t, \dots, t^{n-1}$  then the matrix in this basis for multiplication by  $t$  is shown as below;

$$T = \begin{bmatrix} 0 & & & -a_0 \\ 1 & 0 & & -a_1 \\ & 1 & \ddots & \vdots \\ & & \ddots & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{bmatrix}, a_{n-i} = (-\alpha)^i \binom{n}{i}$$

By using binomial formula, this matrix has polynomial characteristic  $f(t) = (t - \alpha)^n$ .

Therefore by Jordan normal form, it suffices to prove that  $\dim \ker(\alpha I - T) = 1$ .

By rank-nullity theorem, it suffices to prove that  $\text{rk}(\alpha I - T) \geq n - 1$ , for an eigenvalue it has geometric multiplicity at least 1. For  $\alpha = 0$  it is proved and for  $\alpha \neq 0$  the matrix is,

$$\alpha I - T = \begin{bmatrix} \alpha & & & a_0 \\ -1 & \alpha & & a_1 \\ & -1 & \ddots & \vdots \\ & & \ddots & \alpha & a_{n-2} \\ & & & -1 & \alpha + a_{n-1} \end{bmatrix}$$

The above explained matrix is turns into the matrix are shown as below;

$$\begin{bmatrix} 1 & & & \frac{a_0}{\alpha} \\ & 1 & & \frac{a_1}{\alpha} + \frac{a_0}{\alpha^2} \\ & & \ddots & \vdots \\ & & & 1 & \frac{a_{n-2}}{\alpha} + \dots + \frac{a_0}{\alpha^{n-1}} \\ & & & & 1 + \frac{a_{n-1}}{\alpha} + \frac{a_{n-2}}{\alpha^2} + \dots + \frac{a_0}{\alpha^n} \end{bmatrix}$$

Therefore, by Gaussian elimination the rank of matrix is  $\text{rk}(\alpha I - T) \geq n - 1$ .

## 3. a

To determine the matrix of the linear operator,

Suppose  $T$  be the linear operator from the  $R$ -module  $V$  to  $V$ .

That is,  $T: V \rightarrow V$  defined as,

$$T(v) = (t^3 + 3t + 2)v$$

Since, the  $R$ -module  $V$  is the vector space and satisfies that relation,

$$(t^3 + 3t + 2)v = 0$$

And choose a basis for  $V$  as  $\{e_1, e_2, e_3\}$ .

Then,

$$T(e_1) = (t^3 + 3t + 2)(e_1)$$

$$T(e_2) = (t^3 + 3t + 2)(e_2)$$

$$T(e_3) = (t^3 + 3t + 2)(e_3)$$

But, the relation is that,

$$(t^3 + 3t + 2)v = 0$$

Then,

$$T(v) = 0$$

$$(t^3 + 3t + 2)v = 0$$

Then, the vector  $v$  cannot be zero vector for the linear operator.

Then, the vectors  $e_1, e_2, e_3$  cannot be zero.

Then,

$$t^3 + 3t + 2 = 0$$

This characteristic equation is associated with the determinant,

$$t^3 + 3t + 2 = \begin{vmatrix} t & 0 & -1 \\ 2 & t & 0 \\ 3 & -1 & t \end{vmatrix}$$

Hence, the required matrix of the linear operator of multiplication by  $t$  with respect to a basis is

$$\begin{bmatrix} t & 0 & -1 \\ 2 & t & 0 \\ 3 & -1 & t \end{bmatrix}.$$

4. a

Consider the provided statement to prove that  $A = II - B$  is a presentation matrix for the module. As it is provided that  $V$  be a  $F[t]$  module,  $B = (v_1, \dots, v_n)$  is a basis for  $V$ .

It is to be shown that  $\text{im } A = \ker(q : F[t]^n \rightarrow V)$  where the value of  $q$  is given as,

$$(f_i)_{i=1}^n \mapsto \sum_{i=1}^n f_i v_i$$

The inclusion  $\subset$  is clear for  $B$  is multiplication by  $t$  on  $V$ .

For other direction, it is assumed that  $(f_i)_{i=1}^n \in \ker p$  that is,

$$\sum_{i=1}^n f_i v_i = 0, \forall f_i = \sum a_y t^j$$

Then,

$$\begin{aligned} \sum_{i=1}^n f_i v_i &= \sum_{i=1}^n \sum_{j \geq 0} a_y t^j v_i \\ &= \sum_{i=1}^n \left[ a_{i0} v_i + \sum_{j \geq 1} a_y t^{j-1} ((t - B)v_i + Bv_i) \right] \\ &= (q \circ A) \left( \sum_{j \geq 0} a_y t^{j-1} \right)_{i=1}^n + \sum_{i=1}^n \left[ a_{i0} v_i + \sum_{j \geq 1} a_y t^{j-1} Bv_i \right] \\ &= \sum_{i=1}^n \left[ a_{i0} v_i + \sum_{j \geq 1} a_y t^{j-1} Bv_i \right] \end{aligned}$$

As from the above result  $q \circ A = 0$ , then repeat this process and the following result is obtained as shown below;

$$\begin{aligned} \sum_{i=1}^n f_i v_i &= \sum_{i=1}^n \sum_{j \geq 0} a_y B^j v_i \\ &= \sum_{i=1}^n b_i v_i \\ &= 0 \end{aligned}$$

For some coefficients  $b_i \in F$  i.e., for each contribution from the  $f_i$  above could be put in the form of  $(q \circ A)(w)$  for some  $w \in F[t]^n$ . Therefore,  $\text{im } A \supset \ker q$ .

5. a



Consider the provided statement to prove that the characteristic polynomial of the provided matrix is  $f(t)$ .

[Comment](#)

Step 2 of 3 ^

Provided matrix is,  $\begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix}$

As it is known that the Eigen-values is calculated by the characteristic equation of the matrix that is from  $|A - \lambda I| = 0$ . Then,

$$\begin{aligned} \left| \begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| &= 0 \\ \left| \begin{bmatrix} -\lambda & -a_0 \\ 1 & -a_1 - \lambda \end{bmatrix} \right| &= 0 \\ (\lambda) \cdot (-a_1 - \lambda) - (-a_0) &= 0 \\ \lambda^2 - a_1\lambda + a_0 &= 0 \end{aligned}$$

By solving this polynomial, it is finding that the  $\lambda$  has distinct roots that is it has distinct Eigen values is shown as below;

$$\lambda = \frac{1}{2} \left( a_1 - \sqrt{a_1^2 - 4a_0} \right), \lambda = \frac{1}{2} \left( a_1 + \sqrt{a_1^2 - 4a_0} \right)$$

Therefore, the characteristics polynomial is defined as shown as below:

$$f(t) = \left( t - \frac{1}{2} \left( a_1 - \sqrt{a_1^2 - 4a_0} \right) \right) \left( t - \frac{1}{2} \left( a_1 + \sqrt{a_1^2 - 4a_0} \right) \right)$$

When the value of  $t = \frac{1}{2} \left( a_1 - \sqrt{a_1^2 - 4a_0} \right)$  then  $\left( t - \frac{1}{2} \left( a_1 - \sqrt{a_1^2 - 4a_0} \right) \right) = 0$  and when the value of  $t = \frac{1}{2} \left( a_1 + \sqrt{a_1^2 - 4a_0} \right)$  then  $\left( t - \frac{1}{2} \left( a_1 + \sqrt{a_1^2 - 4a_0} \right) \right) = 0$ .

For the characteristics polynomial will be zero, when  $\left( t - \frac{1}{2} \left( a_1 - \sqrt{a_1^2 - 4a_0} \right) \right) = 0$  and

$$\left( t - \frac{1}{2} \left( a_1 + \sqrt{a_1^2 - 4a_0} \right) \right) = 0$$

As it has distinct Eigen-values so the characteristics polynomial of the matrix  $f(t)$  is **proved**.

6. a

Consider the provided statement to classify finitely generated modules over the ring  $\mathbb{C}[\epsilon]$  and it is also given that  $\epsilon^2 = 0$ .

[Comment](#)

Step 2 of 2 ^

As it is known that a module over the ring  $R = \frac{\mathbb{C}[\epsilon]}{(\epsilon^2)}$  has the same concept as a pair  $(M, f)$

where  $M$  is a complex vector space and a function  $f: M \rightarrow M$  is a linear map such that  $f \circ f = 0$ .

To reach  $R$ -module  $M$  can be assign the pair as  $(V, f)$  with  $V = M$  and underlying complex vector of  $M$  and  $f: v \in V \leftrightarrow ev \in V$

For the each pair  $(V, f)$  can be assign  $R$ -module  $M$  that coincides with  $V$  as a complex vector space and the multiplication by  $\epsilon$  is provided as  $\epsilon \cdot v = f(v), \forall v \in M$ .

## Section 9

1. a

(a)

To determine whether or not the modules over  $\mathbb{C}[x, y]$  is free,

Suppose  $\mathbb{C}[x, y]$  be the polynomial ring and  $V$  be the finitely generated module over the polynomial ring  $\mathbb{C}[x, y]$  and the presentation matrix is,

$$A = \begin{bmatrix} x^2 + 1 & x \\ x^2 y + x + y & xy + 1 \end{bmatrix}$$

Then, the finitely generated module  $V$  has two generators  $v_1$  and  $v_2$  and the two relations

$$(x^2 + 1)v_1 + (x^2 y + x + y)v_2 = 0$$

And

$$xv_1 + (xy + 1)v_2 = 0$$

Then, the rank of  $A(c)$  for every point  $c \in \mathbb{C}^2$

Then, the rank of  $A(c)$  for every point  $c \in \mathbb{C}^2$  is,

$$\begin{aligned} m - r &= 2 - 2 \\ &= 0 \end{aligned}$$

Since,  $V$  is a free module of rank  $r$  if and only if the matrix  $A(c)$  has rank  $m - r$  at every point of  $c \in \mathbb{C}^k$

This implies that, the finitely generated module  $V$  is a free module of rank 2.

Hence, it concludes that the finitely generated module modules over  $\mathbb{C}[x, y]$  is free.

(b)

To determine whether or not the modules over  $\mathbb{C}[x, y]$  is free,

Suppose  $\mathbb{C}[x, y]$  be the polynomial ring and  $V$  be the finitely generated module over the polynomial ring  $\mathbb{C}[x, y]$  and the presentation matrix is,

$$A = \begin{bmatrix} xy - 1 \\ x^2 - y^2 \\ y \end{bmatrix}$$

Then, the finitely generated module  $V$  has three generators  $v_1, v_2$  and  $v_3$  and the one relation

$$(xy - 1)v_1 + (x^2 - y^2)v_2 + yv_3 = 0$$

Then, find the rank of  $A(c)$  for every point  $c \in \mathbb{C}^2$ .

Take,  $c \in \mathbb{C}^2$  such that,

$$c = (x, y) = (1, 1)$$

Then, the rank of  $A(c)$  for every point  $(1, 1) = c \in \mathbb{C}^2$  is 1.

But,

$$\begin{aligned} m - r &= 3 - 1 \\ &= 2 \end{aligned}$$

Then, the matrix  $A(c)$  does not have the rank 2 for every point  $c \in \mathbb{C}^2$ .

Since,  $V$  is a free module of rank  $r$  if and only if the matrix  $A(c)$  has rank  $m - r$  at every point of  $c \in \mathbb{C}^k$

This implies that, the finitely generated module  $V$  is not a free module.

Hence, it concludes that the finitely generated module modules over  $\mathbb{C}[x, y]$  is not a free module.

(c)

To determine whether or not the modules over  $\mathbb{C}[x, y]$  is free,

Suppose  $\mathbb{C}[x, y]$  be the polynomial ring and  $V$  be the finitely generated module over the polynomial ring  $\mathbb{C}[x, y]$  and the presentation matrix is,

$$A = \begin{bmatrix} x-1 & x \\ y & y+1 \\ x & y \\ x^2 & 2y \end{bmatrix}$$

Then, the finitely generated module  $V$  has four generators  $v_1, v_2, v_3$  and  $v_4$  and the two relations,

$$(x-1)v_1 + yv_2 + xv_3 + x^2v_4 = 0$$

And

$$xv_1 + (y+1)v_2 + yv_3 + 2yv_4 = 0$$

Then, find the rank of  $A(c)$  for every point  $c \in \mathbb{C}^2$ .

Then, the rank of  $A(c)$  for every point  $c \in \mathbb{C}^2$  is,

$$\begin{aligned} m-r &= 3-2 \\ &= 1 \end{aligned}$$

Since,  $V$  is a free module of rank  $r$  if and only if the matrix  $A(c)$  has rank  $m-r$  at every point of  $\mathbb{C}^k$ .

This implies that, the finitely generated module  $V$  is a free module of rank 2.

Hence, **it concludes that the finitely generated module modules over  $\mathbb{C}[x, y]$  is free.**

## 2. a

Consider the provided statement to prove that module is free.

The provided module is,

$$A = \begin{bmatrix} 1 & x \\ y & x+3 \\ x & y \\ x^2 & y^2 \end{bmatrix}$$

Let it is assumed that there are four generators  $u_1, \dots, u_4$  and there are two relations as

$$u_1 + yu_2 + xu_3 + x^2u_4 = 0 \text{ and } xu_1 + (x+3)u_2 + yu_3 + y^2u_4 = 0.$$

[Comment](#)

Step 2 of 2 ^

The rank of the matrix is shown as below as the order of the matrix is  $4 \times 2$ ,

$$\begin{aligned} \text{Rank}(A) &\leq \min(m, n) \\ &\leq (4, 2) \\ &\leq 2 \end{aligned}$$

As from the theorem 14.9.1, let  $V$  is finitely generated free module over the ring

$R = \mathbb{C}[x_1, \dots, x_k]$ . It is assumed that  $A$  is a presentation matrix for  $V$  and it is the evaluation of  $A$  at a point  $\mathbb{C}^k$  and denoted by  $A(c)$ .

Then  $V$  is a free module of rank  $r$  if and only if the matrix  $A(c)$  has rank  $m-r$  at every point  $c$ . Hence, provided statement is free by exhibiting a basis is **proved**.

## 3. a

To describe the module over the ring  $\mathbb{C}[x, y]$

Suppose  $\mathbb{C}[x, y]$  be the polynomial ring and  $V$  be an abelian group over ring  $\mathbb{C}[x, y]$ .

Then,  $V$  is module over the ring  $\mathbb{C}[x, y]$  if there exists a function  $f : \mathbb{C}[x, y] \times V \rightarrow V$  defined as,

$$f[(r_1, r_2)(v, u)] = (r_1 v, r_2 u)$$

Then, for any  $(r_1, r_2), (r_3, r_4) \in \mathbb{C}[x, y], (v, u), (v_1, u_1), (v_2, u_2) \in V$

Then,

$$\begin{aligned} (r_1, r_2)[(v_1, u_1) + (v_2, u_2)] &= (r_1, r_2)(v_1 + v_2, u_1 + u_2) \\ &= (r_1 v_1 + r_1 v_2, r_2 u_1 + r_2 u_2) \\ &= (r_1, r_2)(v_1, u_1) + (r_1, r_2)(v_2, u_2) \end{aligned}$$

And

$$\begin{aligned} [(r_1, r_2) + (r_3, r_4)](v, u) &= (r_1 + r_3, r_2 + r_4)(v, u) \\ &= (v r_1 + v r_3, u r_2 + u r_4) \\ &= (r_1, r_2)(v, u) + (r_3, r_4)(v, u) \end{aligned}$$

And

$$\begin{aligned} [(r_1, r_2)(r_3, r_4)](v, u) &= (r_1 r_3, r_2 r_4)(v, u) \\ &= (r_1 r_3 v, r_2 r_4 u) \\ &= (r_1(r_3 v), r_2(r_4 u)) \\ &= (r_1, r_2)[(r_3 v), (r_4 u)] \\ &= (r_1, r_2)[(r_3, r_4)(v, u)] \end{aligned}$$

And

$$\begin{aligned} (1, 1)(v, u) &= (1 \cdot v, 1 \cdot u) \\ &= (v, u) \end{aligned}$$

Therefore, it concludes that  $V$  is modules over the ring  $\mathbb{C}[x, y]$  in the terms of complex vector space.

#### 4. a

Consider the provided statement to prove the easy half of the theorem of Quillen and Suslin.

If  $V$  is free then  $V$  is finitely generated free module over the ring  $R = \mathbb{C}[x_1, \dots, x_k]$ . It is assumed that  $A$  is a presentation matrix for  $V$  and it is the evaluation of  $A$  at a point  $\mathbb{C}^k$  and denoted by  $A(c)$ .

[Comment](#)

#### Step 2 of 2 ^

Let  $A_1, A_2, \dots, A_m$  be the column vectors of matrix  $A$ . It is assumed that  $V$  is isomorphic to  $R^k$ . Then, there is an exact sequence of  $R^n \rightarrow R^n \rightarrow R^k \rightarrow 0$ . Let  $X$  is a matrix that denotes the second map such that  $X(X_1, \dots, X_k) = Y_k$ .

On the other hand, it is to be shown that  $R^n$  is the direct sum of the module  $\langle A_1, \dots, A_m \rangle$  and the module generated by  $X_i$ 's. In other words, a matrix  $M$  can be find that in such a way that  $Y_n = (A | X)M$ .

Since, in any evaluation  $v$  the later module is always of rank  $k$ . Then, the matrix  $X(v)$  is of rank  $k$ . Hence, the first module is of rank  $n-k$  then  $A(v)$  has rank  $n-k$ .

Therefore, provided statement is **proved**.

#### 5. a

To prove the module  $V$  is not a free module over the polynomial ring  $R$ ,

Suppose  $R = \mathbb{Z}[\sqrt{-5}]$  be a ring in the imaginary quadratic field  $\mathbb{Q}[\sqrt{-5}]$  and  $V$  be the finitely generated module presented by the presentation matrix  $A$  such that,

$$A = \begin{bmatrix} 2 \\ 1 + \delta \end{bmatrix}$$

Then, the finitely generated module  $V$  has two generators  $v_1$  and  $v_2$  and the one relation,

$$2v_1 + (1 + \delta)v_2 = 0$$

Then, the rank of the presentation matrix  $A$  for every point  $c \in \mathbb{Z}$  is 1.

Since, in the quotient ring  $\frac{R}{P}$ , the ideal  $P$  is prime ideal.

And when  $A$  is in the quotient ring  $\frac{R}{P}$ , then the residue of  $A$  must have the rank 1 as the rank of the presentation matrix  $A$  for every point  $c \in \mathbb{Z}$  is 1.

Then, the residue of  $A$  in the quotient ring  $\frac{R}{P}$  has rank 1 for every prime ideal  $P$ .

Then, the rank of  $A(c)$  for every point  $c \in \mathbb{Z}$  is 1.

Since, the finitely generated module  $V$  is a free module of rank  $r$  if and only if the matrix  $A(c)$  has rank  $m - r$  at every point of  $c \in \mathbb{C}^t$ .

But,

$$\mathbb{Z}[\sqrt{-5}] \neq \mathbb{C}$$

Then, the module  $V$  is not free module over the polynomial ring  $R$ .

Hence, it is proved that the finitely generated module  $V$  is not a free module over the polynomial ring  $R = \mathbb{Z}[\sqrt{-5}]$ .

## Miscellaneous Problem

1. a

Gaussian integers are defined as those integers whose both real and complex parts are defined as the integers.

[Comment](#)

Step 2 of 4 ^

Consider an  $R$ -module structure on the Abelian group  $G$  and consider a ring homomorphism defined as;

$$R \rightarrow \text{End}(G)$$

Since, it is clear that;

$$\text{End}(\mathbb{Z}/5\mathbb{Z}) \cong \mathbb{Z}/5\mathbb{Z}$$

This is a ring.



Now, define a mapping;

$$\phi: \mathbb{Z}[i] \rightarrow \mathbb{Z}/5\mathbb{Z}$$

Then every  $\phi$  must send;

$$1 \rightarrow [1]$$

And the identity element to some element  $x$  of  $\mathbb{Z}/5\mathbb{Z}$  such that;

$$\begin{aligned} x^2 &= -[1] \\ &= [4] \end{aligned}$$

Then there are two choices for  $x$ ;

$$x = [2]$$

Or;

$$\begin{aligned} x &= -[2] \\ &= [3] \end{aligned}$$

Since, both the above obtained choices of the classes defines a ring homomorphism so, define another mapping;

$$f: \mathbb{Z}[X] \rightarrow \mathbb{Z}/5\mathbb{Z}$$

Such that;

$$f(X) = [2] \text{ or } [3]$$

**Hence, the required number of ways is two.**

2. a

Let  $G$  be a finite generated abelian group of order say  $m$ . Then by classification theorem for finitely generated abelian group

$$G \cong \mathbb{Z}^r \oplus \frac{\mathbb{Z}}{m_1\mathbb{Z}} \oplus \frac{\mathbb{Z}}{m_2\mathbb{Z}} \oplus \dots \oplus \frac{\mathbb{Z}}{m_k\mathbb{Z}}, \text{ where the product } m_1 m_2 m_3 \dots m_k \text{ is the prime factorization}$$

of  $m$ , here  $\mathbb{Z}^r$  denotes the torsional free part of  $G$ .

[Comment](#)

Step 2 of 2 ^

Consider  $\mathbb{Z}/(6)$ ,

This ring is same as the ring  $\mathbb{Z}/6\mathbb{Z}$ .

Since prime factorization of 6 is given by

$$6 = 2 \times 3, \text{ where } 2 \text{ and } 3 \text{ both are prime numbers.}$$

Also  $\mathbb{Z}/(6)$  is a cyclic group of order 6 and thus finitely generated abelian group.

Now use the classification theorem for finitely generated groups

So,

$$\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$$

$$\text{Thus, } \mathbb{Z}/(6) \cong \mathbb{Z}/(2) \oplus \mathbb{Z}/(3)$$

**Therefore, all finitely generated modules over the ring  $\mathbb{Z}/(6)$  are classified as**

$$\mathbb{Z}/(6) \cong \mathbb{Z}/(2) \oplus \mathbb{Z}/(3).$$

3. a



Let  $G$  be a finite abelian group. Then by Fundamental theorem of Finite Abelian group which states that every finite abelian group is isomorphic to a direct product of cyclic groups, there exists cyclic groups say  $C_{n_1}, C_{n_2}, C_{n_3}, \dots, C_{n_r}$  such that following holds

$$G \cong C_{n_1} \times C_{n_2} \times C_{n_3} \times \dots \times C_{n_r}, \text{ where order of each cyclic group } C_{n_i} \text{ is given by } n_i, 1 \leq i \leq r$$

Let  $A$  be a finite abelian group.

Then, there exists cyclic groups say  $C_{n_1}, C_{n_2}, C_{n_3}, \dots, C_{n_r}$  such that following holds

$$A \cong C_{n_1} \times C_{n_2} \times C_{n_3} \times \dots \times C_{n_r}, \text{ such that order of each cyclic group } C_{n_i} \text{ is given by } n_i, 1 \leq i \leq r$$

Let  $b_i$  be the generator of the cyclic group  $C_{n_i}$  for every  $i, 1 \leq i \leq r$ .

Also let  $\varphi: A \rightarrow \mathbb{C}^*$  be a non-trivial homomorphism.

So, there exists  $a_o \in A$  such that  $\varphi(a_o) \neq 1$ .

Let  $a \in A$  be an arbitrary element of  $A$ .

$1$  is the identity of the multiplicative group  $\mathbb{C}^*$ .

$$\text{Let } S = \sum_{a \in A} \varphi(a),$$

Assume on contrary that  $S \neq 0$ .

Consider  $\varphi(a_o)S$

Then,

$$\begin{aligned} \varphi(a_o)S &= \varphi(a_o) \left( \sum_{a \in A} \varphi(a) \right) \\ &= \sum_{a \in A} \varphi(a) \varphi(a_o) \\ &= \sum_{a \in A} \varphi(aa_o) \end{aligned}$$

Since  $A \cong C_{n_1} \times C_{n_2} \times C_{n_3} \times \dots \times C_{n_r}$ ,

So for every  $a \in A$

$$a = (b_1^{p_1}, b_2^{p_2}, \dots, b_r^{p_r}), \text{ where } p_i \in \mathbb{Z}, p_i \leq n_i \text{ for every } 1 \leq i \leq r$$

Also since  $a_o \in A$

$$\text{So, } a_o = (b_1^{q_1}, b_2^{q_2}, \dots, b_r^{q_r})$$

Thus,

$$aa_o = (b_1^{p_1+q_1}, b_2^{p_2+q_2}, \dots, b_r^{p_r+q_r})$$

Now if  $p_i + q_i \geq n_i$  for any  $i, 1 \leq i \leq r$  and since  $C_{n_i}$  is cyclic, so  $p_i + q_i \in \{0, 1, 2, 3, \dots, n_i - 1\}$  under modulo  $n_i$  arithmetic.

Hence,  $b_i^{p_i+q_i} \in A$  for every  $i, 1 \leq i \leq r$

Then,

$$aa_o \in A \text{ for every } a \in A$$

Let  $aa_o = x$ , then  $x \in A$  is an arbitrary element of  $A$ .

So,

$$\begin{aligned} \varphi(a_o)S &= \sum_{a \in A} \varphi(aa_o) \\ &= \sum_{x \in A} \varphi(x) \\ &= S \end{aligned}$$

Thus,  $\varphi(a_o)S = S$

This implies that  $(\varphi(a_o) - 1)S = 0$ .

But since  $\varphi(a_o) \neq 1$

Hence  $S = 0$ , this is a contradiction

This contradiction arises due to the wrong assumption that  $S \neq 0$ .

Thus, it can be concluded that  $S = 0$ .

**Therefore, for any finite abelian group  $A$  and a non-trivial homomorphism  $\varphi: A \rightarrow \mathbb{C}^*$  the value of the expression  $\sum_{a \in A} \varphi(a)$  is 0.**

4. a

Let  $A$  be a  $n \times n$  matrix. If  $A$  has distinct  $n$  eigenvalues then  $A$  is said to be diagonalizable. More importantly there exists a matrix  $P$  whose columns are the eigenvectors of the matrix  $A$  such that  $P^{-1}AP$  is a diagonal matrix. And the diagonal entries are the nothing but the eigenvalues of the matrix  $A$

Let  $A$  be a  $2 \times 2$  integer matrix such that it is diagonalizable. Then it has 2 distinct eigenvalues say  $\lambda_1$  and  $\lambda_2$ .

Let  $D = Q^{-1}AP$  such that  $D$  is a diagonal matrix.

Let,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ where } a, b, c, d \in \mathbb{Z}$$

$$Q = \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \text{ where } \begin{pmatrix} e \\ g \end{pmatrix} = x_1, \begin{pmatrix} f \\ h \end{pmatrix} = x_2 \text{ are the eigenvectors corresponding to } \lambda_1 \text{ and } \lambda_2$$

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \text{ where } \lambda_1 \text{ and } \lambda_2 \text{ are the distinct eigenvalues of } A$$

Now since  $\lambda_1$  and  $\lambda_2$  are the distinct eigenvalues of  $A$ ,

So,

$$Ax_1 = \lambda_1 x_1$$

$$Ax_2 = \lambda_2 x_2$$

Since  $D = Q^{-1}AP$ ,

Thus,

$$\begin{aligned} A^{-1}QD &= P \\ &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \\ &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} e\lambda_1 & f\lambda_2 \\ g\lambda_1 & h\lambda_2 \end{pmatrix} \\ A^{-1}QD &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} e\lambda_1 & f\lambda_2 \\ g\lambda_1 & h\lambda_2 \end{pmatrix} \\ &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \left[ \lambda_1 \begin{pmatrix} e \\ g \end{pmatrix} + \lambda_2 \begin{pmatrix} f \\ h \end{pmatrix} \right] \end{aligned}$$

Since  $\{x_1, x_2\}$  are the eigenvectors,

$$\begin{aligned} A^{-1}QD &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \left[ \lambda_1 \begin{pmatrix} e \\ g \end{pmatrix} + \lambda_2 \begin{pmatrix} f \\ h \end{pmatrix} \right] \\ &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} [\lambda_1 x_1 + \lambda_2 x_2] \\ &= A^{-1}(Ax_1 + Ax_2) \\ &= x_1 + x_2 \end{aligned}$$

Hence,

$$\begin{aligned} P &= A^{-1} Q D \\ &= \begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ &= Q \end{aligned}$$

Thus,  $P = Q$

Therefore, when an integer  $2 \times 2$  matrix  $A$  is diagonalized by  $Q^{-1} A P$ , then  $P = Q$ . Also the matrix  $P$  is unique upto rearrangement of the eigenvectors of the matrix  $A$ .

5. a

A lattice is defined as the fundamental algebraic structure which consists of the partially ordered set in which every two elements have a unique supremum and a unique infimum.

[Comment](#)

Step 2 of 4 ^

Consider;

$$\begin{aligned} L: Y &\rightarrow \{\text{lattice in } \mathbf{R}^n\} \\ &= GL_n(\mathbf{R}) / GL_n(\mathbf{Z}) \end{aligned}$$

Such that;

$$g \rightarrow g[\mathbf{Z}^2]; g \in GL_n$$

Now, given an algebraic group  $G$  say,  $GL_n$  and subgroup  $H$  say,  $A = SO_Q$

Here,  $SO_Q$  is large enough such that the quotient group  $SO_Q(\mathbf{R}) / SO_Q(\mathbf{Z})$  is compact for the quotient topology

Now, introduce a representation;

$$\rho: G \rightarrow GL_N$$

Here,  $GL_N$  acts on quadratic forms and a line  $l \subset A^N$  whose stabilizers is  $H$ , then it forms a pair  $(\rho, l)$  such that;

$$G/H \rightarrow \mathbf{P}^{N-1}$$

This is into a projective space.

[Comment](#)

Step 4 of 4 ^

Claim: the solution to Pell's equation for non-square  $d$  so, that it will be easier to generalize it.

Let  $d$  be a non-square integer and;

$$Q(x, y) = x^2 - dy^2$$

So that;

$$SO_Q = \left\{ \begin{bmatrix} a & b \\ db & a \end{bmatrix} \mid a^2 - db = 1 \right\}$$

Therefore, the matrices  $A$  in  $GL_2(\mathbb{R})$  is of the form of  $\boxed{\begin{bmatrix} a & b \\ db & a \end{bmatrix}; a^2 - db = 1}$

6. a

Matrix is defined as the rectangular array which has elements of the form of,  $a_{ij}$ , and are substituted row and column wise

a.

Right multiplication by elements of  $GL_2(\mathbb{Z})$  corresponds to products of elementary row operations on the matrices in  $\mathbb{C}^{2 \times 2}$

After applying reduction formulas on a given matrix which will be of the form of;

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Or;

$$\begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}$$

Or;

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

And;

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Since, every matrix is conjugate that is similar to a unique Jordan form.

Thus, the orbits are generated as follows;

$$\left( \coprod_{\lambda \in \mathbb{C}} GL_2(\mathbb{C}) \times \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \right) \coprod \left( \coprod_{\lambda_1 \in \mathbb{C}, \lambda_2 \in \mathbb{C}} GL_2(\mathbb{C}) \times \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \right)$$

b.

Consider the linear system;

$$AX = b$$

Here;

$$A \in M_{m,n}(\mathbb{R})$$

$$b \in \mathbb{R}^m$$

$$x \in \mathbb{R}^n$$

Then there exists an invertible matrices  $P \in M_m(\mathbb{R})$  and  $Q \in M_n(\mathbb{R})$  such that after the invertible change of variables;

$$\begin{aligned} z &= (z^1, \dots, z^n)^r \\ &= Qx \end{aligned}$$

Then the system becomes;

$$\begin{aligned} z^i &= p^i \forall i = 1, \dots, r \\ 0 &= p^i \forall i = r+1, \dots, m \end{aligned}$$

Where,

$$\begin{aligned} P &= (p^1, \dots, p^m) \\ &= AP \end{aligned}$$

Here,  $r$  is the rank of  $A$

Where;

$$p^1 = \{d_1, 0, 0, \dots\}$$

$$p^2 = \{a_2, d_2, 0, \dots\}$$

$$p^3 = \{a_3, b_3, d_3, 0, \dots\}$$

And so on

Hence, the invertible matrix  $AP$  has the following hermitian normal form;

$$\begin{bmatrix} d_1 & 0 & 0 & 0 & \dots \\ a_2 & d_2 & 0 & 0 & \\ a_3 & b_3 & d_3 & 0 & \\ \vdots & \vdots & & & \ddots \end{bmatrix}$$

Where, the entries are non-negative  $a_2 < d_2, a_3, b_3 < d_3$  and so on

7. a

A subring is defined as the subset of a ring  $R$  which itself is a ring and satisfies binary operations on addition and shares the same multiplicative identity as  $R$

[Comment](#)

Step 2 of 4 ^

Consider  $R = \mathbb{Q}[t]$

Let  $s, t \in R$

Since,  $s, t$  are integral over  $R$

Thus;

$$R[s, t] \subseteq R$$

In particular  $st$  and  $s-t$  are both in  $R$  so,  $R$  is a ring extension then  $S$  will be defined as the integral closure of  $R$  in  $S$

Now, consider the ring  $\mathbb{Z}$

Here,  $\mathbb{Z}$  is integrally closed in its field of fractions

Claim: Every  $x/y \in \mathbb{Q}$  which is integral over  $\mathbb{Z}$  is actually in  $\mathbb{Z}$

Let;

$$x/y \in \mathbb{Z}$$

Such that;

$$(x, y) = 1$$

And, suppose that  $x/y$  is a root of;

$$a^n - c_{n-1}a^{n-1} - \dots - c_1a - c_0$$

Substitute  $x/y$  and multiplying by  $y^{n-1}$ , then;

$$x^n/y = c_{n-1}x^{n-1} + c_{n-2}x^{n-2}y + \dots + c_0y^{n-1}$$

Here;  $c_{n-1}x^{n-1} + c_{n-2}x^{n-2}y + \dots + c_0y^{n-1}$  is an integer

Hence, the  $x^n/y$  will also be an integer

Thus,  $y$  divides  $x^n$

Since,  $(x, y) = 1$ , this implies that;

$$y = 1$$

And;

$$x/y \in \mathbb{Z}$$

Also, since the unique factorisation domain is integrally closed in its field of fractions.

Since,  $\mathbb{Z}$  is integrally closed in  $\mathbb{Q}$

However,  $\mathbb{Z}$  is not integrally closed in  $\mathbb{C}$

**Therefore,  $R$  is finitely generated  $S$ -module.**

## 8. a

A group  $G$  is said to be finitely generated if there exists finitely many elements  $x_1, x_2, \dots, x_n \in G$  such that every  $x \in G$  can be written in the below mentioned form

$$x = m_1 x_1 + m_2 x_2 + \dots + m_n x_n, \text{ where } m_i \in \mathbb{Z}, 1 \leq i \leq n$$

A number  $\alpha$  is said to be algebraic if there exists a monic polynomial  $f(x)$  such that  $f(\alpha) = 0$ .

**(a)**

Let  $\alpha \in \mathbb{C}$  be an algebraic integer and let  $q$  be the degree of its minimal polynomial.

Then,

The minimal polynomial of  $\alpha$  has integer coefficients, that is, the coefficients lie in  $\mathbb{Z}$ .

Since for all  $u \geq q$

$$\alpha^u = a_1 + a_2 \alpha + \dots + a_{q-1} \alpha^{q-1}, \text{ where } a_i \in \mathbb{Z}, 1 \leq i \leq q-1$$

Hence,

$$\mathbb{Z}[\alpha] = \mathbb{Z} \oplus \mathbb{Z}\alpha \oplus \mathbb{Z}\alpha^2 \oplus \dots \oplus \mathbb{Z}\alpha^{q-1}$$

Thus it can be concluded that  $\{1, \alpha, \alpha^2, \dots, \alpha^{q-1}\}$  generate  $\mathbb{Z}[\alpha]$  as an abelian group.

Conversely, assume that  $\mathbb{Z}[\alpha]$  is finitely generated, with generators say  $a_1, a_2, \dots, a_q$ , where

$$a_i = f_i(\alpha), 1 \leq i \leq q \text{ for some polynomial } f_i \in \mathbb{Z}[x].$$

Let  $n \in \mathbb{Z}$  such that  $n > \deg f_i$  for  $i = 1, 2, 3, \dots, q$ .

Now,

$$\alpha^n = \sum_{j=1}^q b_j a_j, \quad b_j \in \mathbb{Z}$$

Then clearly,

$$\alpha^n - \sum_{j=1}^q b_j f_j(\alpha) = 0$$

Choose the polynomial as

$$f(X) = X^n - \sum_{j=1}^q b_j f_j(X)$$

Clearly all the coefficients of the above chosen polynomial lie in  $\mathbb{Z}$ .

Since,  $n > \deg f_i$  for  $i = 1, 2, 3, \dots, q$

Thus the polynomial  $f(X) \in \mathbb{Z}[X]$  is monic.

Also evaluate  $f(X)$  at  $\alpha$ .

So,

$$\begin{aligned} f(\alpha) &= \alpha^n - \sum_{j=1}^q b_j f_j(\alpha) \\ &= 0 \end{aligned}$$

Hence,  $f(\alpha) = 0$

So,  $\alpha$  is an algebraic integer.

**Therefore, for any complex number  $\alpha$  and  $\mathbb{Z}[\alpha]$  be the subring of  $\mathbb{C}$  generated by  $\alpha$ , the complex number  $\alpha$  is an algebraic integer if and only if  $\mathbb{Z}[\alpha]$  is a finitely generated abelian group.**



(b)

Let  $\alpha, \beta$  be algebraic integers.

Then,  $\mathbb{Z}[\alpha], \mathbb{Z}[\beta]$  are finitely generated abelian groups.

Let  $f(x), g(x)$  be the monic polynomials corresponding to algebraic integers  $\alpha, \beta$  respectively such that

Degree  $f(x) = m$  and Degree  $g(x) = n$

Let  $\{1, \alpha, \alpha^2, \dots, \alpha^{m-2}, \alpha^{m-1}\}$  be the generator set of  $\mathbb{Z}[\alpha]$  and  $\{1, \beta, \beta^2, \dots, \beta^{n-2}, \beta^{n-1}\}$  be the generator set of  $\mathbb{Z}[\beta]$ .

Then,  $\{\alpha^i \beta^j \mid 0 \leq i \leq m-1, 0 \leq j \leq n-1\}$  forms the generator set of  $\mathbb{Z}[\alpha, \beta]$ .

Thus  $\mathbb{Z}[\alpha, \beta]$  is finitely generated. And since  $\mathbb{Z}[\alpha], \mathbb{Z}[\beta]$  are abelian so is  $\mathbb{Z}[\alpha, \beta]$ .

**Hence, if  $\alpha$  and  $\beta$  are algebraic integers, then the subring  $\mathbb{Z}[\alpha, \beta]$  of  $\mathbb{C}$  that they generate is a finitely generated abelian group.**

(c)

Let  $W$  denote the set of all algebraic integers of the field  $\mathbb{C}$ .

Let  $\alpha, \beta \in W$ .

Thus  $\alpha$  and  $\beta$  are algebraic integers.

Then,  $\mathbb{Z}[\alpha], \mathbb{Z}[\beta]$  are finitely generated abelian groups

Use the result

If  $\alpha$  and  $\beta$  are algebraic integers, then the subring  $\mathbb{Z}[\alpha, \beta]$  of  $\mathbb{C}$  that they generate is a finitely generated abelian group.

Since,  $\mathbb{Z}[\alpha, \beta]$  is finitely generated abelian group.

Also,  $\mathbb{Z}[\alpha, \beta]$  is a ring and since a ring is closed under addition and multiplication.

So,  $\alpha + \beta$  and  $\alpha - \beta$  belong to  $\mathbb{Z}[\alpha, \beta]$  and  $\alpha\beta$  also belongs to  $\mathbb{Z}[\alpha, \beta]$ .

That is,

$$\alpha + \beta \in \mathbb{Z}[\alpha, \beta]$$

$$\alpha - \beta \in \mathbb{Z}[\alpha, \beta]$$

$$\alpha\beta \in \mathbb{Z}[\alpha, \beta]$$

Also  $\mathbb{Z}[\alpha + \beta], \mathbb{Z}[\alpha - \beta]$  and  $\mathbb{Z}[\alpha\beta]$  are subgroups of  $\mathbb{Z}[\alpha, \beta]$ .

So,

$\mathbb{Z}[\alpha + \beta], \mathbb{Z}[\alpha - \beta]$  and  $\mathbb{Z}[\alpha\beta]$  are all finitely generated being subgroups of finitely generated abelian group.

Hence,  $\mathbb{Z}[\alpha + \beta], \mathbb{Z}[\alpha - \beta]$  and  $\mathbb{Z}[\alpha\beta]$  are all finitely generated abelian groups.

Use the result

For any complex number  $\alpha$  and  $\mathbb{Z}[\alpha]$  be the subring of  $\mathbb{C}$  generated by  $\alpha$ , the complex number  $\alpha$  is an algebraic integer if and only if  $\mathbb{Z}[\alpha]$  is a finitely generated abelian group.

Thus  $\alpha + \beta, \alpha - \beta$  and  $\alpha\beta$  are all algebraic integers.

**Hence,  $W$  the set of all algebraic integers of the field  $\mathbb{C}$  forms a ring.**

9. a

a.

Consider  $L$  be a lattice the define;

$$L^* = \{w / (v, w) \in \mathbb{Z} \forall v \in L\}$$

Suppose  $\{v_1, \dots, v_k\}$  be linearly dependent vectors

Or is enough to show that for each  $w$  in  $v$  can be expressed in more than one way as a linear combination of  $v_1, \dots, v_k$

Since,  $v_i$  is linearly dependent then there exists  $a_1, \dots, a_k \in \mathbb{R}$  not all zero such that;

$$c_1 v_1 + \dots + c_k v_k = 0$$

Since, for given  $w$  there are  $b_1, \dots, b_k \in \mathbb{R}$

So, consider the linear combination;

$$(a_1 + b_1) v_1 + \dots + (a_k + b_k) v_k = 0 + w \\ = w$$

This constitutes a different linear combination then;  $b_1 v_1 + \dots + b_k v_k$  because not all of the  $c_i$  are zero.

**Thus,  $L$  has a basis  $B = (v_1, \dots, v_k)$  a set of  $k$  vectors that spans  $L$**

b.

[Comment](#)

Step 4 of 6 ^

Suppose  $B$  is a basis for the given lattice, that is;

$$L^* = L(B)$$

This implies that, it will also be a basis for the vector space, that is  $\text{span}(B)$

Now, define;

$$L(B) = \{Bx : x \in \mathbb{Z}^n\}$$

This can be further extended to the basis matrices  $B$  in which the columns are linearly dependent.

Here, clearly  $B$  is a matrix which is having rational entries, then  $L(B)$  is defined as a lattice.

Thus, the basis  $L^*$  can be derived from the basis  $B$

**Therefore,  $L^*$  can be defined as;**

$$L^* = L(B)$$

**Where,  $B$  is the basis.**

c.

Every lattice is the complete join of all its one-element sublattices

If  $L$  is a lattice and;

$$|L^*| = \alpha$$

This is a regular infinite cardinal

Now, suppose that  $|L^*|$  is uncountable and;

$$L^* = X \cup Y$$

Then either;

$$|X| = |L^*|$$

Or;

$$|Y| = |L^*|$$

**So, if there is a lattice, then it must be joining irreducible in the sublattice  $L$**

d.

Consider  $L \subset L^* \subset \mathbb{Z}^2$

And, let  $L^*$  has index  $m$  in  $\mathbb{Z}^2$

Suppose  $L$  is a sublattice of  $\mathbb{Z}^2$  of index  $p^m$ ,  $p$  is prime

This implies that  $L$  corresponds to an element of  $S_{p^m}$

For finding the index first suppose that  $L$  is not contained in  $p\mathbb{Z}^2$

Then, the image of  $L$  will be;

$$\mathbb{Z}^2 / p\mathbb{Z}^2 = (\mathbb{Z}/p\mathbb{Z})^2$$

This will be order  $p^2$

This means that;

$$L^* = p\mathbb{Z}^2 + L$$

Then;

$$[L^* : \mathbb{Z}^2] = p$$

Therefore;

$$[L^* : L] = p^{m-1}$$

10. a

(a)

To prove that multiplicative group  $\mathbb{Q}^*$  of a rational number is isomorphic to the direct sum of a cyclic group whose order is provided.

Let  $\langle -1 \rangle = \{-1, 1\}$  with the obvious group structure making it isomorphic to  $\frac{\mathbb{Z}}{2\mathbb{Z}}$ .

Let  $\langle p \rangle = \{p^k \mid k \in \mathbb{Z}\}$  with the obvious group structure making it isomorphic to  $\mathbb{Z}$ . Then,

$$\langle -1 \rangle \oplus \langle p \rangle \rightarrow \mathbb{Q}^*, \sigma \oplus p^k p \mapsto \prod_{p \text{ prime}} p^k p$$

It is noted that  $\langle p \rangle$  is a free abelian group. This map is a group homomorphism because

$$(\sigma, k_1, \dots) \cdot (\tau, l_1, \dots) \text{ get mapped to } \sigma\tau \prod p^{k_i + l_i} = \sigma \prod p^{k_i} \tau \prod p^{l_i}.$$

This map is surjective since any rational number  $\frac{r}{s}$  has prime decompositions for  $r$  and  $s$ .

Similarly, this is injective since something on the right is equal to 1 if and only if  $\sigma = 1$  and  $k_p = 0, \forall p$ .

(b)

Now it is to be proving that the additive group  $\mathbb{Q}^+$  of rational numbers is not a direct sum of two proper subgroups.

Let  $G, H$  be the proper non-trivial subgroup of  $\mathbb{Q}^+$  and from the proposition 2.11.4 it suffices to show that they have non-trivial intersection.

If  $\frac{p}{q} \in G, \frac{s}{t} \in H$  for some  $p, s \neq 0$  then,

$$\begin{aligned} qs \left( \frac{p}{q} \right) &= pt \left( \frac{s}{t} \right) \\ &= ps \in G \cap H \end{aligned}$$

While it is assumed that  $ps \neq 0$ , therefore provided condition is **proved**.

(c)

Now it is to be proved that the quotient group  $\frac{\mathbb{Q}^+}{\mathbb{Z}^+}$  is not a direct sum of cyclic groups.

It is assumed that for  $H_k$  cyclic  $\frac{\mathbb{Q}^+}{\mathbb{Z}^+} \approx \bigoplus_k H_k$ . Let the image of  $\frac{p_k}{q_k} + \mathbb{Z}^+ \in H_k$  is generate  $H_k$  for  $(p_k, q_k) = 1$ . Let  $sp_k + tq_k = 1$  for some integers  $s, t$  then,

$$\begin{aligned} \frac{1}{q_k} + \mathbb{Z}^+ &= \frac{sp_k}{q_k} + \frac{tq_k}{q_k} + \mathbb{Z}^+ \\ &= s \frac{p_k}{q_k} + \mathbb{Z}^+ \in H_k \end{aligned}$$

Therefore, it can be assumed that  $\frac{1}{q_k}$  generates  $H_k$ .

Now it is consider that the element  $\frac{1}{q_1^2} + \mathbb{Z}^+ \in \frac{\mathbb{Q}^+}{\mathbb{Z}^+}$ .  $\frac{1}{q_1^2} + \mathbb{Z}^+$  Has a decomposition is shown as below;

$$\frac{1}{q_1^2} + \mathbb{Z}^+ = \sum_k \frac{r_k}{q_k} + \mathbb{Z}^+$$

Where  $r_k \mid q_k$  for all but finitely many values of  $k$ , adding it to itself  $q_1$  times; then

$$\frac{1}{q_1^2} + \mathbb{Z}^+ = \sum_k \frac{q_1 r_k}{q_k} + \mathbb{Z}^+$$

Since the direct sum decomposition,  $q_k \mid q_1 r_k, \forall k \neq 1$  and  $q_1 \mid q_1 r_1 - 1$ . This property implies that  $q_1 \mid 1$  so that there is contradiction. Hence, provided statement is proved.

## Chapter 15

### Section 1

1. a

Field is defined as equivalent to the ring where its non-zero elements form an abelian group under multiplication.

[Comment](#)

Step 2 of 4 ^

Let  $R$  be an integral domain that contains a field  $F$  as subring and that is finite dimensional when viewed as vector space over  $F$ .

To prove:  $R$  is a field

Consider that each non-zero element  $\alpha$  of  $R$  is algebraic, and so;

$$F[\alpha] \subseteq R$$

This is a field.

Hence, each nonzero element  $R$  has an inverse then it will show that  $R$  is a field

Claim: each nonzero element  $\alpha$  is algebraic

Since, the dimensions is a finite number  $n$ , the  $n+1$  vectors  $1, \alpha, \alpha^2, \dots, \alpha^n$  must be linearly independent over  $F$ .

That is there are elements  $a_0, a_1, \dots, a_n$  of  $F$ , not all 0, such that;

$$a_0 + a_1\alpha + a_2\alpha^2 + \dots + a_n\alpha^n = 0$$

Let  $j$  be the smallest index such that  $a_j \neq 0$

Then the polynomial equation becomes;

$$a_j\alpha^j + a_{j+1}\alpha^{j+1} + \dots + a_n\alpha^n = 0$$

Factoring out  $\alpha^j$ ;

$$\alpha^j (a_j + a_{j+1}\alpha + \dots + a_n\alpha^{n-j}) = 0$$

Since,

$$\alpha \neq 0$$

And, since  $R$  is an integral domain where;

$$\alpha^j \neq 0$$

Again since  $R$  is an integral domain and since the product;

$$\alpha^j (a_j + a_{j+1}\alpha + \dots + a_n\alpha^{n-j}) = 0$$

The factor;

$$a_j + a_{j+1}\alpha + \dots + a_n\alpha^{n-j} = 0$$

Hence,

$$\alpha (a_{j+1} + a_{j+2}\alpha + \dots + a_n\alpha^{n-j-1}) = -a_j$$

Since;

$$a_j \neq 0$$

And, since  $F$  is a field, the inverse of  $\alpha$  is;

$$(-a_j)^{-1} (a_{j+1} + a_{j+2}\alpha + \dots + a_n\alpha^{n-j-1})$$

**Therefore,  $R$  is a field**

## 2. a

Discriminant is defined as the coefficients of the polynomial equation whose values gives the information about the roots of the polynomial.

[Comment](#)

Step 2 of 4 ^

Let  $F$  be a field, not of characteristic 2, and let;

$$x^2 + bx + c = 0$$

This be a quadratic equation with coefficients into prove: that  $\delta$  is an element of  $F$  such that;

$$\delta^2 = b^2 - 4c, \quad x = \frac{-b + \delta}{2} \text{ solves the quadratic equation in } F \text{ also, prove that if the } b^2 - 4c$$

discriminant is not a square, the polynomial has no root in  $F$

Let  $\alpha$  and  $\beta$  be roots such that;

$$x^2 + bx + c = (x - \alpha)(x - \beta)$$

Then the discriminant  $\delta$  is defined as  $\delta^2$ , where;

$$\delta = (\beta - \alpha)$$

Now;

$$\alpha + \beta = -b$$

And;

$$\alpha\beta = c$$

Therefore;

$$\delta^2 = \alpha^2 + \beta^2$$



Further, put  $x = (-b + \delta)/2$  in the quadratic equation  $x^2 + bx + c = 0$ :

$$\begin{aligned} \left(\frac{-b+\delta}{2}\right)^2 + b\left(\frac{-b+\delta}{2}\right) + c &= \frac{(-b+\delta)^2}{4} + \frac{-b^2 + \delta b}{2} + c \\ &= \frac{b^2 + \delta^2 - 2b\delta - 2b^2 + 2\delta b + 4c}{4} \\ &= \frac{\delta^2 - b^2 + 4c}{4} \end{aligned}$$

This is contained in  $F$  as this is of the form of rational numbers

Thus,  $x = (-b + \delta)/2$  solves the quadratic equation in  $F$

[Comment](#)

Step 4 of 4 ^

Now, if the discriminant is not a square then, the term;

$$b^2 - 4ac$$

This will not be a rational number.

That is this will be an irrational number

And, irrationals does not lie in field

**Therefore, if the  $b^2 - 4ac$  discriminant is not a square, the polynomial has no root in  $F$**

3. a

**Solution:** We will determine all the subfields of  $\mathbb{C}$  that are dense subsets of  $\mathbb{C}$ .

Now notice that if  $K$  is a subfield of  $\mathbb{C}$  such that  $K \subseteq \mathbb{R}$ , then clearly  $K$  can not be dense in  $\mathbb{C}$ .

So we will only consider the subfields  $K$  of  $\mathbb{C}$  such that  $K \not\subseteq \mathbb{R}$ .

Since  $\mathbb{Q}$  is the smallest subfield of  $\mathbb{C}$ ,  $K$  must contain  $\mathbb{Q}$ . Now by our assumption of  $K$  we have  $K \not\subseteq \mathbb{R}$ , therefore there exists an element  $\alpha \in K$  such that  $\alpha \notin \mathbb{R}$ .

It follows that  $\mathbb{Q}(\alpha)$  is contained in  $K$ .

**Claim:** The above defined  $\mathbb{Q}(\alpha)$  is dense in  $\mathbb{C}$ .

**Proof of the Claim:** Notice that the set  $\{1, \alpha\}$  forms an  $\mathbb{R}$ -basis of  $\mathbb{C}$ .

Then any element  $z$  in  $\mathbb{C}$  can be written as

$$z = a + \alpha b, \quad \text{where } a, b \in \mathbb{R}. \quad (1)$$

Let us now consider an arbitrary open set  $A$  in  $\mathbb{C}$ .

In order to show that  $\mathbb{Q}(\alpha)$  is dense in  $\mathbb{C}$ , it is enough to show that

$$\mathbb{Q}(\alpha) \cap A \neq \emptyset.$$

Now by (1) every element of  $A$  can be written in that form. So for our simplicity let us assume

$$A := \{a + \alpha b \mid a \in A_1, b \in A_2\}$$

where  $A_1$  and  $A_2$  are some intervals containing  $a$  and  $b$  in  $\mathbb{R}$  respectively.

Now recall that  $\mathbb{Q}$  is a dense subset of  $\mathbb{R}$ . Then by the definition of dense we have

$$A_1 \cap \mathbb{Q} \neq \emptyset \quad \text{and} \quad A_2 \cap \mathbb{Q} \neq \emptyset.$$

Therefore there exist rationals  $x$  and  $y$  in  $A_1$  and  $A_2$  respectively such that

$$x + \alpha y \in A.$$

This follows that

$$x + \alpha y \in A \cap \mathbb{Q}(\alpha).$$

This follows that

$$A \cap \mathbb{Q}(\alpha) \neq \emptyset.$$

Since  $A$  is an arbitrary open set in  $\mathbb{C}$ , it yields that every open set in  $\mathbb{C}$  intersects  $\mathbb{Q}(\alpha)$  non-trivially.

Then by the definition it follows that  $\mathbb{Q}(\alpha)$  is dense in  $\mathbb{C}$ .

This completes the proof of the Claim.

Now recall that  $\mathbb{Q}(\alpha) \subset K$ .

Since  $\mathbb{Q}(\alpha)$  is dense in  $\mathbb{C}$ ,  $K$  is also dense in  $\mathbb{C}$ .

Therefore every subfields of  $\mathbb{C}$  that are not contained in  $\mathbb{R}$  are dense in  $\mathbb{C}$ .

This completes the proof.

## Result

3 of 3

Every subfields of  $\mathbb{C}$  that are not contained in  $\mathbb{R}$  are dense in  $\mathbb{C}$

## Section 2

### 1. a

Polynomial is defined as an expression of more than two terms mainly referred as the sum of the terms of different powers.

[Comment](#)

### Step 2 of 3 ^

Let  $\alpha$  be a complex root of the polynomial  $x^3 - 3x + 4$

To find: the inverse of  $\alpha^2 + \alpha + 1$  in the form of  $a + b\alpha + c\alpha^2$  with  $a, b, c$  in  $\mathbb{Q}$

For the proof first suppose;

$$1 = (a + b\alpha + c\alpha^2)(1 + \alpha + \alpha^2)$$

This can be written as;

$$1 = a + (a + b)\alpha + (a + b + c)\alpha^2 + (b + c)\alpha^3 + c\alpha^4$$

Now, note that;

$$\alpha^3 = 3\alpha - 4$$

And,

$$\alpha^4 = 3\alpha^2 - 4\alpha$$

So, the equation is as follows;

$$\begin{aligned} 1 &= a + (a + b)\alpha + (a + b + c)\alpha^2 + (b + c)(3\alpha - 4) + c(3\alpha^2 - 4\alpha) \\ &= (a - 4b - 4c) + (a + 4b - c)\alpha + (a + b + 4c)\alpha^2 \end{aligned}$$

Since,  $1, \alpha$  and  $\alpha^2$  are linearly independent over  $\mathbb{Q}$ , then a system of linear equations with solution is obtained;

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & -4 & -4 \\ 1 & 4 & -1 \\ 1 & 1 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$= \frac{1}{49} \begin{pmatrix} 17 & 12 & 20 \\ -5 & 8 & -3 \\ -3 & -5 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$= \frac{1}{49} \begin{pmatrix} 17 \\ -5 \\ -3 \end{pmatrix}$$

And hence,

$$(1 + \alpha + \alpha^2)^{-1} = \frac{1}{49}(17 - 5\alpha - 3\alpha^2)$$

Therefore, the inverse of  $\alpha^2 + \alpha + 1$  in the form of  $a + b\alpha + c\alpha^2$  with  $a, b, c$  in  $\mathbb{Q}$  is;

$$\boxed{\frac{1}{49}(17 - 5\alpha - 3\alpha^2)}$$

## 2. a

An irreducible polynomial is defined as a non-constant polynomial that cannot be factored into the product of two non-constant polynomials.

[Comment](#)

Step 2 of 3 ^

Let  $f(x) = x^n - a_{n-1}x^{n-1} + \dots \pm a_0$  be the irreducible polynomial over  $F$ , and let  $\alpha$  be a root of  $f$  in an extension field  $K$

To determine: the element  $\alpha^{-1}$  explicitly in terms of  $\alpha$  and of the coefficients  $a_i$

For the proof consider  $\alpha$  be the root of  $f$  over  $\mathbb{F}_q$

Then;

$$[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$$

Hence;

$$\mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$$

And, thus;

$$\alpha \in \mathbb{F}_{q^m}$$

Claim: If  $\alpha^{-1} \in \mathbb{F}_{q^n}$  is a root of  $f$

Then,  $(\alpha^{-1})^q$  is also a root of  $f$

Now, write;

$$f = x^n - a_{n-1}x^{n-1} + \dots \pm a_0$$

Then;

$$\begin{aligned} f((\alpha^{-1})^q) &= x^n - a_{n-1}x^{n-1} + \dots \pm a_0 \\ &= x^{qn} - a_{n-1}^q x^{q(n-1)} + \dots \pm a_0^q \\ &= (x^n - a_{n-1}x^{n-1} + \dots \pm a_0)^q \\ &= f(\alpha^{-1})^q \end{aligned}$$

Thus,  $\alpha, \alpha^q, \dots, \alpha^{q(n-1)}$  are roots of  $f$

Thus;

$$\boxed{f(\alpha^{-1})^q = \alpha^n - a_{n-1}\alpha^{q(n-1)} + \dots \pm a_0}$$

3. a

**Given:** Let  $\beta = \omega \sqrt[3]{2}$ , where  $\omega = e^{\frac{2\pi i}{3}}$  and considering  $K = \mathbb{Q}(\beta)$ .

**To Prove:** The equation  $x_1^2 + x_2^2 + \dots + x_k^2 = -1$  has no solution  $x_i$  in  $K$ , for  $k \geq 1$ .

**Proof:** Let us consider the cubic polynomial  $f(x) = x^3 - 2$ .

Notice that both  $\sqrt[3]{2}$  and  $\beta$  are roots of  $f(x)$ .

It follows that  $f(x)$  is the irreducible polynomial for both  $\sqrt[3]{2}$  and  $\beta$  over  $\mathbb{Q}$ . Let us now consider the following lemma.

**Lemma:** Let  $F$  be a field and  $\alpha$  and  $\beta$  be two elements of the field extensions  $K/F$  and  $L/F$ . Suppose that  $\alpha$  and  $\beta$  are algebraic over  $F$ . There is an isomorphism of fields  $\sigma : F(\alpha) \rightarrow F(\beta)$  that is the identity on  $F$  and that sends  $\alpha \rightarrow \beta$  if and only if the irreducible polynomials for  $\alpha$  and  $\beta$  over  $F$  are equal.

Now it follows from the above Lemma that  $\mathbb{Q}(\sqrt[3]{2})$  and  $\mathbb{Q}(\beta)$  are isomorphic.

Let us now consider the polynomial

$$g(x_1, x_2, \dots, x_k) = x_1^2 + x_2^2 + \dots + x_k^2 + 1.$$

Now note that  $\mathbb{Q}(\sqrt[3]{2})$  is a subfield of  $\mathbb{R}$ . Therefore the polynomial  $g(x_1, x_2, \dots, x_k)$  has no solution in  $\mathbb{R}$ , for every  $k \geq 1$ .

This completes the proof.

## Result

The result follows from the isomorphic condition of two fields  $\mathbb{Q}(\beta)$  and  $\mathbb{Q}(\sqrt[3]{2})$ .

## Section 3

1. a

Let  $F$  be a field and let  $\alpha$  be an element that generates a field extension of  $F$  of degree 5.

To prove: That  $\alpha^2$  generates the same extension

For the proof consider  $F$  to be a subfield of a field  $K$ , denote the dimension of  $K$  as an  $F$ -vector space by  $[K : F]$

It is already known that;

$$[K : F] = p$$

Where,  $p$  is a prime number

Then there are no fields properly between  $F$  and  $K$

Now, consider;

$$[F[\alpha] : F] = 5$$

And, so  $F[\alpha^2]$  must be either  $F$  or  $F[\alpha]$

Claim: That  $F[\alpha^2]$  cannot be  $F$  and so, it must equal  $F[\alpha]$

If;

$$F[\alpha^2] = F$$

Then;

$$\alpha^2 \in F$$

So, that  $\alpha$  satisfies the degree 2 polynomial;

$$x^2 - \alpha^2 \in F[x]$$

That is;

$$[F[\alpha] : F] \text{ does not have value } 5$$

**Therefore,  $\alpha^2$  generates the same extension**

## 2. a

**To Prove:** The polynomial  $x^4 + 3x + 3$  is irreducible over  $\mathbb{Q}(\sqrt[3]{2})$ .

**Proof:** Let us consider  $f(x) = x^4 + 3x + 3$ .

Firstly we will use **Eisenstein criterion** to show that  $f(x)$  is irreducible over  $\mathbb{Q}$ .

Recall the **Eisenstein criterion** as follows:

Suppose we have the following polynomial with integer coefficients.

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$

If there exists a prime number  $p$  such that the following three conditions all apply:

(1)  $p$  divides each  $a_i$ , for  $i \neq n$ .

(2)  $p$  does not divide  $a_n$ , and

(3)  $p^2$  does not divide  $a_0$

then  $p$  is irreducible over the rational numbers.

In case of our  $f(x)$  let us take  $p = 3$ .

Then by **Eisenstein criterion**  $f(x)$  is irreducible over  $\mathbb{Q}$ .

Let us assume that  $\alpha$  is a root of  $f(x)$ .

Then  $\alpha$  defines a field extension of degree 4 over  $\mathbb{Q}$ .

Now notice that

$$\begin{aligned} [\mathbb{Q}(\alpha, \sqrt[3]{2}) : \mathbb{Q}] &= [\mathbb{Q}(\alpha, \sqrt[3]{2}) : \mathbb{Q}(\sqrt[3]{2})][\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] \\ &= 4 \times 3 \\ &= 12. \end{aligned}$$

Therefore the extension  $\mathbb{Q}(\alpha, \sqrt[3]{2})$  over  $\mathbb{Q}$  has degree 12.  
 The irreducible polynomial for  $\alpha$  over  $\mathbb{Q}(\sqrt[3]{2})$  must have degree 4.  
 It follows that  $f(x)$  is a irreducible polynomial over  $\mathbb{Q}(\sqrt[3]{2})$ .  
 This completes the proof.

## Result

Using Eisenstein Criterion we have proved that  $x^4 + 3x + 3$  is irreducible over  $\mathbb{Q}(\sqrt[3]{2})$ .

## 3. a

A splitting field of a polynomial with coefficients in a field is a smallest field extension of that field over which the polynomial gets factored.

[Comment](#)

Step 2 of 3 ^

Let  $\zeta_n = e^{2\pi i/n}$

To prove:  $\zeta_5 \notin \mathbb{Q}(\zeta_7)$

Since, the field  $\mathbb{Q}(\zeta_n)$  is the splitting field for the polynomial  $x^n - 1$

Since, 7 is prime

So, the degree will be;

$$[\mathbb{Q}(\zeta_7) : \mathbb{Q}] = 7 - 1 = 6$$

The minimal polynomial for  $\zeta_7$  is the 7-cyclotomic polynomial, that is;

$$\begin{aligned} \phi_7(x) &= \sum_{n=0}^6 x^n \\ &= \prod_{n=1}^6 (x - \zeta^n) \end{aligned}$$

Since, the polynomial is irreducible, then there exist  $\sigma \in G$  such that;

$$\sigma(\zeta) = \zeta^3$$

Then;

$$\sigma^2(\zeta) = \zeta^2$$

$$\sigma^3(\zeta) = \zeta^6$$

$$\sigma^4(\zeta) = \zeta^4$$

$$\sigma^5(\zeta) = \zeta^5$$

$$\sigma^6(\zeta) = \zeta$$

Hence, from the above explanation it is clear that  $\zeta_5 \notin \mathbb{Q}(\zeta_7)$

## 4. a



An irreducible polynomial is defined as the non-constant polynomial that cannot be factored into the product of two non-constant polynomials.

[Comment](#)

Step 2 of 7 ^

Let  $\zeta_n = e^{2\pi i/n}$

To determine: The irreducible polynomial over  $\mathbb{Q}$  and over  $\mathbb{Q}(\zeta_3)$

a.

To determine the irreducibility of  $\zeta_4$

Here,  $\zeta_n$  satisfies  $x^n - 1$  that is;

$$(x^n - 1) / (x - 1)$$

For  $\zeta_2 = -1$  the minimal polynomial is  $x + 1$

b.

To determine the irreducibility of  $\zeta_6$

And, for  $\zeta_6 = -\zeta_3$ , then;

$$((-x) - 1) / ((-x) - 1) = x^2 - x + 1$$

[Comment](#)

Step 4 of 7 ^

c.

To determine the irreducibility of  $\zeta_8$

And for  $\zeta_8$ ;

$$\zeta_8^4 = -1$$

So,  $x^4 + 1$  has no roots and no quadratic factors.

d.

To determine the irreducibility of  $\zeta_9$

For  $\zeta_9$ ;

$$\zeta_9^3 = \zeta_3$$

So;

$$(x^3)^2 + x^3 + 1 = x^6 + x^3 + 1$$

This is a polynomial satisfied by it.

Replace  $x$  by  $x + 1$  then Einstein this with  $p = 3$ ;

e.

To determine the irreducibility of  $\zeta_{10}$

Now, for  $\zeta_{10}$ :

$$\zeta_{10}^5 = -1$$

So;

$$x^5 + 1$$

This is a polynomial satisfied by it.

Now, factor;

$$x^5 + 1 = (x + 1)(x^4 - x^3 + x^2 - x + 1)$$

Here,  $\zeta_{10}$  does not have a degree 2.

---

[Comment](#)

---

Step 7 of 7 ^

f.

To determine the irreducibility of  $\zeta_{12}$

Finally, for  $\zeta_{12}$ :

$$x^6 + 1$$

This polynomial factor the sum of cubes to get  $x^4 - x^2 + 1$

Thus, like the above parts  $\zeta_{12}$  does not have a degree 2.

5. a

Degree is defined as the exponential quantity defined on the variable of the given expression.

---

[Comment](#)

---

Step 2 of 4 ^

To determine: The values of  $n$  such that  $\zeta_n$  has degree at most 3 over  $\mathbb{Q}$

For determining the result suppose, if  $p$  is a prime dividing  $n$ , then;

$$\zeta_p^n = 1$$

Hence,

$$\mathbb{Q}(\zeta_n) \supset \mathbb{Q}(\zeta_p)$$

So, by using the theorem which states that;

Let  $p$  be a prime. The cyclotomic polynomial  $\phi(x) = x^{p-1} + x^{p-2} + \dots + x + 1$  is irreducible over  $\mathbb{Q}$ .

Then by above result;

$$|\mathbb{Q}(\zeta_p) : \mathbb{Q}| = p - 1$$

So;

$$|\mathbb{Q}(\zeta_n) : \mathbb{Q}| \leq 3$$

This modulus implies that if  $p \mid n$ , then  $p \in \{2, 3\}$

Thus,

$$n = 2^i 3^j$$

Claim: For  $i \leq 2, j \leq 1$ ,

$$|\mathbb{Q}(\zeta_2) : \mathbb{Q}| = 1$$

And,

$$|\mathbb{Q}(\zeta_4) : \mathbb{Q}| = 2$$

Since, the minimal polynomial of  $\zeta_4$  is  $x^2 + 1$ .

Similarly,

$$|\mathbb{Q}(\zeta_3) : \mathbb{Q}| = 2$$

Now, by using the theorem which states that;

Let  $F \subset K \subset L$  be fields. Then,;

$$[L : F] = [L : K][K : F]$$

Therefore both,

$$[L : K] \text{ and } [K : F] \text{ divide } [L : F]$$

But any extension of  $\mathbb{Q}(\zeta_4)$  and  $\mathbb{Q}(\zeta_3)$  will have degree  $\geq 4$  over  $\mathbb{Q}$

Therefore, from the above theorem the claim is proved that  $i \leq 2, j \leq 1$

Now, from the above explanation;

$$n \in \{1, 2, 3, 4, 6, 12\}$$

Also;

$$|\mathbb{Q}(\zeta_n) : \mathbb{Q}| \leq 3 \text{ for } n \in \{1, 2, 3, 4\}$$

Hence, it remains to show that whether  $n \in \{6, 12\}$  are also possible values of  $n$

That is;

$$\zeta_6 = -\zeta_3$$

Hence,

$$\mathbb{Q}(\zeta_6) = \mathbb{Q}(\zeta_3)$$

On the other hand, by using the theorem which states that;

Let  $F \subset K \subset L$  be fields. Then,;

$$[L : F] = [L : K][K : F]$$

Therefore both,

$$[L : K] \text{ and } [K : F] \text{ divide } [L : F]$$

That is;

$$\mathbb{Q}(\zeta_{12}) \supset \mathbb{Q}(\zeta_3, \zeta_4)$$

Which has degree greater than and equals to 6 over  $\mathbb{Q}$  and hence,  $n \neq 12$

**Thus,**  $n \in \{1, 2, 3, 4, 5\}$

6. a

Rational numbers are in the form of a rational number that is  $\frac{p}{q}$  where both  $p, q$  are integers and  $q$  is never equals to zero.

[Comment](#)

Step 2 of 4 ^

Let  $a$  be a positive rational number that is not a square in  $\mathbb{Q}$

To prove: That  $\sqrt[4]{a}$  has degree 4 over  $\mathbb{Q}$

Here, the degree is at most 4.

Now;

$$\begin{aligned}\mathbb{Q} &\subset \mathbb{Q}[\sqrt{a}] \\ &\subset \mathbb{Q}[\sqrt[4]{a}]\end{aligned}$$

Where the former extension is to have a degree of 2

Now to show:  $[\mathbb{Q}[\sqrt[4]{a}]:\mathbb{Q}[\sqrt{a}]] \neq 1$

Now, take;

$$\begin{aligned}\alpha &= \sqrt[4]{a} \\ &= a + b\sqrt{a}; a, b \in \mathbb{Q}\end{aligned}$$

Then;

$$\sqrt{a} = a^2 + ab^2 + 2ab\sqrt{a}$$

That is;

$$\begin{aligned}\sqrt{a}(1 - 2ab) &= a^2 + ab^2 \\ &\in \mathbb{Q}\end{aligned}$$

Since,  $\sqrt{a}$  is irrational so;

$$2ab = 1$$

And;

$$a^2 + ab^2 = 0$$

Then;

$$a = -b^2$$

Simultaneously;

$$ab = \frac{1}{2}$$

That is;

$$\begin{aligned}-b^3 &= \frac{1}{2} \\ &\in \mathbb{Q}\end{aligned}$$

Since, this cannot happen.

Thus;  $[\mathbb{Q}[\sqrt[4]{a}]:\mathbb{Q}[\sqrt{a}]] \neq 1$

**Therefore,  $\sqrt[4]{a}$  has degree 4 over  $\mathbb{Q}$**

7. a

a.

To show: That whether  $i$  is the field  $\mathbb{Q}(\sqrt[4]{-2})$

Suppose;

$$i \in \mathbb{Q}(\sqrt[4]{-2})$$

Then, there exists  $a, b \in \mathbb{Q}$  such that;

$$(a + b\sqrt[4]{-2})^2 = -1$$

But then;

$$\begin{aligned} (a + b\sqrt[4]{-2})^2 &= a^2 + 2ab\sqrt[4]{-2} + b^2\sqrt[4]{-2} \\ &= -1 \end{aligned}$$

Now, by using the proposition which states that;

Let  $\alpha$  be an algebraic element over  $F$ , and let  $f(x)$  be the irreducible polynomial for  $\alpha$  over  $F$ .

If  $f(x)$  has degree  $n$ , that is if  $\alpha$  has degree  $n$  over  $F$

And since;

$$\{1, \sqrt[4]{-2}, \sqrt[4]{-2}^2\}$$

It is linearly independent over  $\mathbb{Q}$ , that is  $a^2 = -1$

**This implies that it is impossible since  $a \in \mathbb{Q}$**

b.

To show: That whether  $\sqrt[3]{5}$  is the field of  $\mathbb{Q}(\sqrt[3]{2})$

$$\text{If } \sqrt[3]{5} \in \mathbb{Q}(\sqrt[3]{2})$$

Then;

$$\begin{aligned} \alpha &= \sqrt[3]{2} + \sqrt[3]{5} \\ &\in \mathbb{Q}(\sqrt[3]{2}) \end{aligned}$$

This must have degree at most 3 over  $\mathbb{Q}$

Now, let;

$$f(x) = x^9 - 21x^6 - 123x^3 - 343$$

Then;

$$f(\alpha) = 0$$

But  $f(x)$  is irreducible by the Eisenstein criterion which states that;

Let  $f(x) = a_n x^n + \dots + a_0$  be an integer polynomial and let  $p$  be a prime integer. Suppose that the coefficients of  $f$  satisfy the following conditions;

First is  $p$  does not divide  $a_n$ ;

Second is  $p$  divides all other coefficients  $a_{n-1}, \dots, a_0$ ;

And, third is  $p^2$  does not divide  $a_0$

That is, since,

$$3 \mid 21, 123$$

While;

$$9 \nmid 343$$

Hence,  $f$  is irreducible and so  $\alpha$  has degree 9 over  $\mathbb{Q}$ , this is a contradiction.

8. a

An algebraic number is any complex number that is a root of a non-zero polynomial in one variable with rational coefficients.

[Comment](#)

Step 2 of 3 ^

Let  $\alpha$  and  $\beta$  be complex numbers.

Consider that if  $\alpha + \beta$  and  $\alpha\beta$  are algebraic numbers,

To prove:  $\alpha$  and  $\beta$  are algebraic numbers.

If both  $\alpha + \beta$  and  $\alpha\beta$  are algebraic, then the extension;

$$K = F(\alpha + \beta, \alpha\beta)$$

This is algebraic over  $F$  by following the definition of the algebraic numbers

Claim: That  $F(\alpha, \beta)/K$  is algebraic, and then the required result follows

Now, consider the polynomial;

$$x^2 - (\alpha + \beta)x + \alpha\beta$$

This has coefficient in  $K$

That means  $\alpha$  and  $\beta$  are both the roots of  $K$

**Therefore,  $\alpha$  and  $\beta$  are algebraic numbers.**

9. a

Irreducible polynomials are defined as those expressions which cannot be factored into the product of two non-constant polynomials.

[Comment](#)

Step 2 of 3 ^

Consider  $\alpha$  and  $\beta$  be complex roots of irreducible polynomials  $f(x)$  and  $g(x)$  in  $\mathbb{Q}[x]$ . Also, let;

$$K = \mathbb{Q}(\alpha)$$

$$L = \mathbb{Q}(\beta)$$

To prove:  $f(x)$  is irreducible in  $L[x]$  if and only if  $g(x)$  is irreducible in  $K[x]$

For the proof consider;

$$\deg f = n$$

$$\deg g = m$$

Now, by multiplicative property of the degree it can be obtained that:

$$|\mathbb{Q}(\alpha, \beta) : \mathbb{Q}| = |\mathbb{Q}(\beta) : K| |K : \mathbb{Q}|$$

Hence, from above it can be said that  $f(x)$  will be irreducible in  $L[x]$  which satisfies the condition:

$$|\mathbb{Q}(\alpha, \beta) : \mathbb{Q}| = mn$$

And, this is possible only if  $g(x)$  is irreducible in  $K[x]$

Therefore,  $f(x)$  is irreducible in  $L[x]$  if and only if  $g(x)$  is irreducible in  $K[x]$

10. a



A field is a complete group consisting of nonzero commutative division ring whose nonzero elements forms an abelian group under multiplication.

[Comment](#)

Step 2 of 4 ^

Consider a field extension  $K/F$  which is an algebraic extension if every element of  $K$  is algebraic over  $F$ . Also, let  $K/F$  and  $L/K$  be algebraic field extensions.

To prove:  $L/F$  is an algebraic extension

For the proof, consider  $\beta \in L$

It is enough to show that  $\beta$  is algebraic over  $F$

For this, let;

$$\beta = \sum_{j=1}^n a_j \beta_j$$

Where,

$$a_j \in K$$

$$\beta_j \in L$$

These are linearly independent and algebraic over  $K$

Again, consider;

$$\beta_j = \sum_{i=1}^m b_i \alpha_i$$

Where,

$$b_i \in F$$

$$\alpha_i \in K$$

These are linearly independent and algebraic over  $K$

Thus,

$$\beta \in F(\alpha_1, \dots, \alpha_m)(\beta_1, \dots, \beta_m)$$

Now, consider the theorem which states that;

Let  $F \subset K \subset L$  be fields. Then,

$$[L:F] = [L:K][K:F]$$

Therefore, both  $[L:K]$  and  $[K:F]$  divide  $[L:F]$

Hence, by using the above theorem;

$$[F(\alpha_1, \dots, \alpha_m)(\beta_1, \dots, \beta_m) : F(\alpha_1, \dots, \alpha_m)] [F(\alpha_1, \dots, \alpha_m) : F] \geq [F(\beta) : F] \quad \text{This contradicts the}$$

$$= \infty$$

fact that a field extension  $K$  that is generated over  $F$  by finitely many algebraic elements is a finite extension. A finite extension is generated by finitely many elements.

Thus,  $\beta$  is algebraic over  $F$

Therefore,  $L/F$  is an algebraic extension

## Section 4

1. a

Consider  $K = \mathbb{Q}(\alpha)$

Where,  $\alpha$  is a root of  $x^3 - x - 1$

To determine: The irreducible polynomial of  $\gamma = 1 + \alpha^2$  over  $\mathbb{Q}$

First,  $\gamma$  satisfies

$$(x-1)(x-2)^2 - 1 = x^3 - 5x^2 + 8x - 5 = 0$$

Now, by theorem which states that;

Let  $F \subset K \subset L$  be fields. Then,;

$$[L:F] = [L:K][K:F]$$

Therefore both,

$$[L:K] \text{ and } [K:F] \text{ divide } [L:F]$$

This theorem implies that;

$$[K:\mathbb{Q}] = [K:\mathbb{Q}(\gamma)][\mathbb{Q}(\gamma):\mathbb{Q}] = 3$$

And so;

$$[\mathbb{Q}(\gamma):\mathbb{Q}] = 3$$

Since,  $\gamma \in \mathbb{Q}$

**Thus,  $x^3 - 5x^2 + 8x - 5$  is the irreducible polynomial of  $\gamma$  over  $\mathbb{Q}$**

## 2. a

Irreducible polynomial is defined as non-constant polynomial that cannot be factored into the product of two non-constant polynomials.

**a**

To determine: The irreducible polynomial over the field  $\mathbb{Q}$

Let;

$$f(x) = (x^2 - 8)^2 - 60 = x^4 - 16x^2 + 4$$

Then;

$$f(\alpha) = 0$$

Claim:  $f$  is irreducible

Here,  $f$  cannot have linear factors since a root must be an integer dividing 4 by the rational root test which states that;

A constraint on rational solutions of a polynomial equation;

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 = 0$$

With integer coefficients when written as a fraction  $x = \frac{p}{q}$  in lowest terms satisfies  $p$  is an integer factor of the constant term  $a_0$  and  $q$  is an integer of the leading coefficient  $a_n$

From the above result these are not roots.

Now, suppose  $f$  had quadratic factors, then by proposition which states that;

Let  $f$  be an integer polynomial with positive leading coefficient. Then  $f$  is an irreducible element of  $\mathbb{Z}[x]$  if and only if it is either a prime integer or a primitive polynomial that is irreducible in  $\mathbb{Q}[x]$

So for some  $a, b, c, d \in \mathbb{Z}$ ;

$$\begin{aligned} f(x) &= (x^2 + ax + b)(x^2 + cx + d) \\ &= x^4 + (a+c)x^3 + (b+ac+d)x^2 + (ad+bc)x + bd \end{aligned}$$

Now;

$$a + c = 0$$

Implies that;

$$\begin{aligned} ad + bc &= a(d - b) \\ &= 0 \end{aligned}$$

Since,  $a \neq 0$  otherwise;

$$b + d = -16$$

And;

$$bd = 4$$

This is a contradiction. So,

$$\begin{aligned} d &= b \\ &= 2 \end{aligned}$$

But then;

$$2 + ac + 2 = -16$$

This implies that;

$$ac = -20$$

This contradicts the fact;

$$a + c = 0$$

**Hence,  $f$  is the irreducible polynomial for  $\alpha$  over  $\mathbb{Q}$**

**b.**

To determine: The irreducible polynomial over the field  $\mathbb{Q}(\sqrt{5})$

Let;

$$\begin{aligned} f(x) &= (x - \sqrt{5})^2 - 3 \\ &= x^2 - 2\sqrt{5}x + 2 \end{aligned}$$

Then;

$$f(\alpha) = 0$$

Here,  $f$  splits if and only if the two roots  $\sqrt{5} \pm \sqrt{3}$  are in  $\mathbb{Q}(\sqrt{5})$

But;

$$\sqrt{3} \notin \mathbb{Q}(\sqrt{5})$$

For otherwise;

$$\sqrt{3} = a + b\sqrt{5}$$

Implies;

$$3 = a^2 + 2ab\sqrt{5} + 5b^2$$

And, further solving for  $\sqrt{5}$  implies that  $\sqrt{5} \in \mathbb{Q}$

This is a contradiction.

**Hence,  $f$  is the irreducible polynomial for  $\alpha$  over  $\mathbb{Q}(\sqrt{5})$**

c.

To determine: The irreducible polynomial over the field  $\mathbb{Q}(\sqrt{10})$

Let;

$$f(x) = x^4 - 16x^2 + 4$$

Now, by using the corollary which states that;

Let  $\mathcal{K}$  be an extension of a field of  $F$ , let  $K$  and  $F'$  be the subfields of  $\mathcal{K}$  that are finite extensions of  $F$  and let  $K'$  denote the subfield of  $\mathcal{K}$  generated by the two fields  $K$  and  $F'$  together. Let;

$$[K' : F] = N$$

$$[K : F] = m$$

$$[F' : F] = n$$

Then  $m$  and  $n$  divide  $N$  and  $N \leq mn$

And also by using the first part above it must have degree either 4 or 8 over  $\mathbb{Q}$

Hence, by the multiplicative property of the degree,  $\sqrt{10}$  has 2 or 1 over  $\mathbb{Q}(\alpha)$

To show:  $\sqrt{10}$  does not have degree 1 over  $\mathbb{Q}(\alpha)$

Now, it remains to show that it is not in  $\mathbb{Q}(\alpha)$

Hence, by lemma this states that;

An element  $\alpha$  of a field extension  $K$  has degree 1 over  $F$  if and only if  $\alpha$  is an element of  $F$

But this is clear since if;

$$\sqrt{10} = a + b\alpha; a, b \in \mathbb{Q}$$

Then;

$$\begin{aligned} 10 &= a^2 + 2ab\alpha + b^2\alpha^2 \\ &= a^2 + 8b^2 + 2ab\sqrt{3} + 2ab\sqrt{5} + 2b^2\sqrt{15} \end{aligned}$$

Implies  $b = 0$  contradicting that  $\sqrt{10} \notin \mathbb{Q}$

Now,  $\alpha$  has degree 4 over  $\mathbb{Q}(\sqrt{10})$  since;

$$\begin{aligned} [\mathbb{Q}(\alpha, \sqrt{10}) : \mathbb{Q}(\sqrt{10})][\mathbb{Q}(\sqrt{10}) : \mathbb{Q}] &= [\mathbb{Q}(\alpha, \sqrt{10}) : \mathbb{Q}(\sqrt{10})] \times 2 \\ &= 8 \end{aligned}$$

And, thus  $f$  is irreducible polynomial for  $\alpha$  over  $\mathbb{Q}(\sqrt{10})$

d.

To determine: The irreducible polynomial over the field  $\mathbb{Q}(\sqrt{15})$

Let;

$$f(x) = x^3 - 8 - 2\sqrt{15}$$

Then;

$$f(\alpha) = 0$$

Now;

$$\begin{aligned} 4 &= [\mathbb{Q}(\alpha, \sqrt{15}) : \mathbb{Q}] \\ &= [\mathbb{Q}(\alpha, \sqrt{15}) : \mathbb{Q}(\sqrt{15})][\mathbb{Q}(\alpha, \sqrt{15}) : \mathbb{Q}] \end{aligned}$$

Then by the first part solved above since;

$$\begin{aligned} \sqrt{15} &= \frac{(\alpha^2 - 8)}{2} \\ &\in \mathbb{Q}(\alpha) \end{aligned}$$

Hence,  $\alpha$  has degree 2 over  $\mathbb{Q}(\sqrt{15})$  and so,  $f$  is the irreducible polynomial over  $\mathbb{Q}(\sqrt{15})$

3. a

Consider the polynomial  $f(x) = x^3 - 2$

To determine: the irreducible polynomial for  $\gamma = \alpha_1 + \alpha_2$  over  $\mathbb{Q}$

Now,  $x^3 - 2$  has solution;

$$\alpha = \sqrt[3]{2}$$

This is clearly a root of  $x^3 - 2$

Then after factoring and applying quadratic formula, the factors will be as follows;

$$x^3 - 2 = (x - \alpha)(x - \alpha\gamma)(x - \alpha\gamma^2)$$

Here,  $\zeta$  is a complex cube root of unity

Now, further  $\zeta$  will become;

$$\gamma^2 + \gamma + 1 = 0$$

And  $\gamma \notin \mathbb{R}$

Hence,  $\gamma \notin \mathbb{Q}(\alpha)$

Now, consider the roots of  $x^3 - 2$  as;

$$\alpha_i; i = 1, 2, 3$$

Now, each of the roots  $\alpha_i; i = 1, 2, 3$  has degree 3 over  $F$

Also, the nine monomials  $\alpha_i^j \alpha_2^k$  with  $0 \leq i, j < 3$  span  $K$  over  $F$

Since, these monomials are not independent that is,  $f$  has a root  $\alpha_1$  in the field  $L$ , it factors in  $L[x]$ , that is;

$$f(x) = (x - \alpha_1)q(x)$$

Then  $\alpha_2$  is a root of  $q(x)$ , so  $\alpha_2$  has degree at most 2 over  $L$

Then the set  $(1, \alpha_2)$  is a basis for  $K$  over the field  $L$  then there exist  $\gamma$  such that;

$$\gamma = \alpha_1^i \alpha_2^j; 0 \leq i < 3, 0 \leq j < 2$$

Therefore, the irreducible polynomial for  $\gamma = \alpha_1 + \alpha_2$  over  $\mathbb{Q}$  is;

$$\boxed{\gamma = \alpha_1^i \alpha_2^j; 0 \leq i < 3, 0 \leq j < 2}$$

## Section 5

1. a

**Solution:** We will express  $\cos(15^\circ)$  in terms of real square roots.

Recall that for any two angles  $\alpha$  and  $\beta$  we have

$$\cos(\alpha - \beta) = \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta).$$

Then note that

$$\begin{aligned} \cos(15^\circ) &= \cos(45^\circ - 30^\circ) \\ &= \cos(45^\circ) \cos(30^\circ) + \sin(45^\circ) \sin(30^\circ) \\ &= \left( \frac{1}{\sqrt{2}} \times \frac{\sqrt{3}}{2} \right) + \left( \frac{1}{\sqrt{2}} \times \frac{1}{2} \right) \\ &= \frac{\sqrt{3}}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} \\ &= \frac{\sqrt{6} + \sqrt{2}}{4}. \end{aligned}$$

Here we are done.

## Result

The value of  $\cos 15^\circ$  is  $\frac{\sqrt{6} + \sqrt{2}}{4}$ .

## 2. a

Polynomial is defines as an expression of more than two algebraic terms, where it is written as the sum of the several terms having different powers.

a.

To find: By the field theory

Consider;

$$\alpha = \frac{2\pi}{5}$$

And;

$$\begin{aligned} z &= e^{i\alpha} \\ &= \cos(\alpha) + i \sin(\alpha) \\ &= x + iy \end{aligned}$$

Now;

$$\begin{aligned} z^5 &= (x + iy)^5 \\ &= 1 \end{aligned}$$

And, using;

$$y^2 = 1 - x^2$$

Then the following is obtained that  $\cos \alpha$  is a root of the polynomial, that is;

$$16x^5 - 20x^3 + 5x - 1$$

Clearly 1 is another root and it can be further factorized

That is;

$$\begin{aligned} 16x^5 - 20x^3 + 5x - 1 &= (x - 1)(16x^4 + 16x^3 - 4x^2 - 4x + 1) \\ &= (x - 1)(4x^2 + 2x - 1)^2 \end{aligned}$$

So, that by the quadratic equation and the fact that;

$$\alpha = \frac{-1 + \sqrt{5}}{4}$$

Hence,  $\alpha$  is constructible over  $\mathbb{Q}$



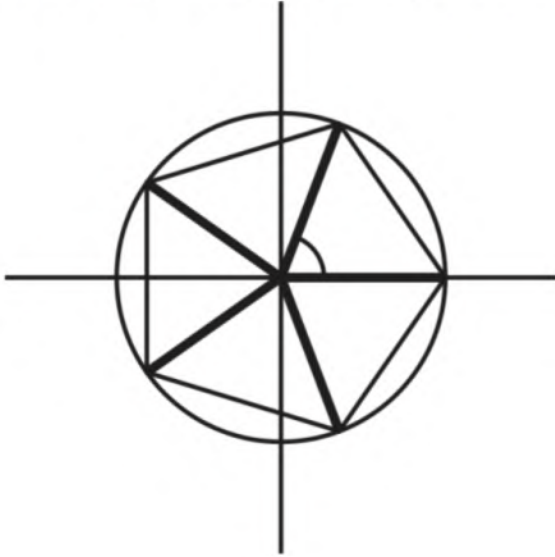
b.

To find: By an explicit construction

Now, to explicitly construct the regular pentagon, start from the points  $(0,0)$  and  $(0,1)$  and build successively  $2\pi/5$  angles, every time constructing the points;

$$p_n = (\cos(2\pi/5), \sin(2\pi/5)), n = 1, \dots, 4$$

And, taking the half-lines through the origin and each  $p_n$ , the obtained figure is dawning below;



Take the circle of radius 1 centered at the origin and construct the points at the intersection of the circle and the half-lines. Connecting these yield the regular pentagon.

3. a

**Solution:** We will propose to prove that the regular 9-gon is not constructible by ruler and compass.

Now note that the angle of a 9-gon is  $40^\circ$ , since  $\frac{360^\circ}{9} = 40^\circ$ .

So if the above one is constructed, then clearly  $\cos 40^\circ$  can be constructed.

Now notice that

$$\begin{aligned} -\frac{1}{2} &= \cos(120^\circ) \\ &= \cos(40^\circ + 80^\circ) \\ &= \cos(40^\circ) \cos(80^\circ) - \sin(40^\circ) \sin(80^\circ) \\ &= \cos(40^\circ)(2 \cos^2(40^\circ) - 1) - \sin(40^\circ)(2 \sin(40^\circ) \cos(40^\circ)) \\ &= \cos(40^\circ)(2 \cos^2(40^\circ) - 1) - 2 \sin^2(40^\circ) \cos(40^\circ) \\ &= \cos(40^\circ)(2 \cos^2(40^\circ) - 1) - 2(1 - \cos^2(40^\circ)) \cos(40^\circ) \\ &= x(2x^2 - 1) - 2x(1 - x^2), \text{ taking } \cos(40^\circ) = x \\ &= 4x^3 - 3x. \end{aligned}$$

Therefore we have

$$\begin{aligned} 4x^3 - 3x + \frac{1}{2} &= 0 \\ \implies 8x^3 - 6x + 1 &= 0. \end{aligned}$$

Now we will prove that the polynomial  $8x^3 - 6x + 1$  is irreducible over  $\mathbb{Q}$ .

Let us consider

$$f(x) = 8x^3 - 6x + 1.$$

Now we have

$$\begin{aligned} f(x+1) &= 8(x+1)^3 - 6(x+1) + 1 \\ &= 8x^3 + 24x^2 + 18x + 3. \end{aligned}$$

By **Eisenstein's Criterion** taking  $p = 3$ ,  $f(x+1)$  is irreducible over  $\mathbb{Q}$ .

It follows that  $f(x)$  is irreducible over  $\mathbb{Q}$ .

Therefore the degree of  $\cos(40^\circ)$  is  $3 \neq 2^n$ . It follows that  $\cos(40^\circ)$  is not constructible.

Hence the regular 9-gon is not constructible by ruler and compass.

Now recall the **Eisenstein's criterion**, which states as follows:

Suppose we have the following polynomial with integer coefficients.

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$

If there exists a prime number  $p$  such that the following three conditions all apply:

- (1)  $p$  divides each  $a_i$ ,  $i \neq n$ .
- (2)  $p$  does not divide  $a_n$ .
- (3)  $p^2$  does not divide  $a_0$ .

Then  $f$  is irreducible over  $\mathbb{Q}$ .

## Result

The regular 9-gon is not constructible by ruler and compass.

## 4. a

Let  $\Delta$  denote the given triangle

The area of the triangle is defined below:

$$A = \frac{1}{2}bh$$

Where;

$A$  = area

$l$  = length

$b$  = breadth

To show: Is it possible to construct a square whose area is equal to that of a given triangle.

Now, construct the length  $s$  such that;

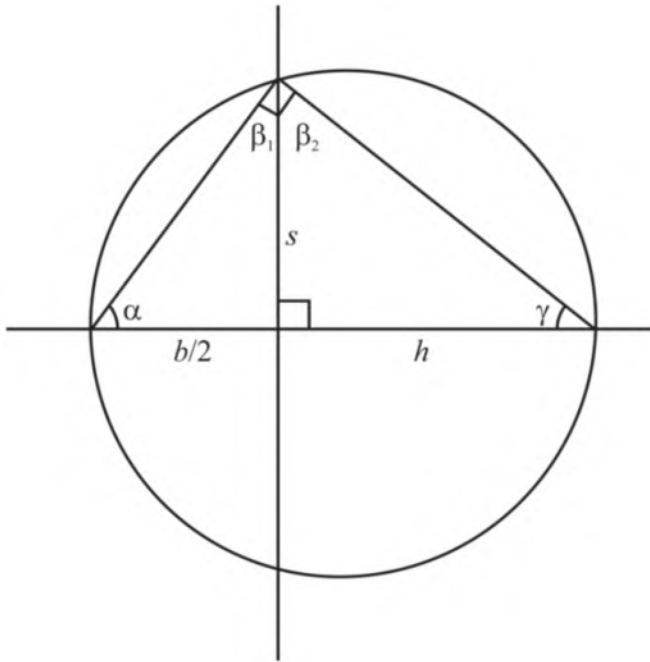
$$s^2 = \frac{1}{2}bh$$

Now, mark the lengths  $\frac{b}{2}$  and  $h$  onto a constructed line so that the two segments are adjacent.

Then take the circle passing through the endpoints of diameter;

$$\frac{b}{2} + h$$

And, then the triangle inscribed in this circle with edges the endpoints, the obtained figure is given below;



Now, check that the height  $s$  of this triangle is exactly the length that is;

$$\frac{s}{h} = \frac{b/2}{s}$$

Denote the angles  $\alpha, \beta, \gamma$  where;

$$\beta = \frac{\pi}{2}$$

This breaks down into;

$$\beta = \beta_1 + \beta_2$$

For the triangle on the left and the triangle on the right, then following can be obtained;

$$\begin{aligned} \frac{s}{h} &= \tan \gamma \\ &= \tan \left( \frac{\pi}{2} - \beta_2 \right) \\ &= \tan (\beta_1) \\ &= \frac{b/2}{s} \end{aligned}$$

5. a

Discriminant is defined as the coefficients of the polynomial equation whose value gives information about the roots of the polynomial.

[Comment](#)

#### Step 2 of 5 ^

Suppose that the determinant is negative

To determine: The line that appears at the end of the proof geometrically

The line of intersection between any two lines is obtained by solving two linear equations with coefficients in  $F$

To find the intersection of a line and the circle use the equation of the line to eliminate one variable from the equation of the circle

This quadratic equation has solutions in the field;

$$F' = F(\sqrt{D}) \text{ Here, } D \text{ is the discriminant}$$

Now, consider the case when the discriminant is negative that is take the case when  $D$  is negative.

For this case there will be no real solution to the equations. Then the line and the circle do not intersect each other;

Consider the intersection between two circles as given below;

$$(x - a_1)^2 + (y - b_1)^2 = c_1$$

$$(x - a_2)^2 + (y - b_2)^2 = c_2$$

[Comment](#)

#### Step 4 of 5 ^

Now, solving the equation separately;

$$(x - a_1)^2 + (y - b_1)^2 = c_1$$

$$(x)^2 - 2(x)(a_1) + (a_1)^2 + (y)^2 - 2(y)(b_1) + (b_1)^2 = c_1$$

$$x^2 - 2xa_1 + a_1^2 + y^2 - 2yb_1 + b_1^2 = c_1$$

Further solving for the second equation;

$$(x - a_2)^2 + (y - b_2)^2 = c_2$$

$$(x)^2 - 2(x)(a_2) + (a_2)^2 + (y)^2 - 2(y)(b_2) + (b_2)^2 = c_2$$

$$x^2 - 2xa_2 + a_2^2 + y^2 - 2yb_2 + b_2^2 = c_2$$

Now, subtract both above solved equations;

$$(x^2 - 2xa_1 + a_1^2 + y^2 - 2yb_1 + b_1^2) - (x^2 - 2xa_2 + a_2^2 + y^2 - 2yb_2 + b_2^2) = c_1 - c_2$$

$$x^2 - 2xa_1 + a_1^2 + y^2 - 2yb_1 + b_1^2 - x^2 + 2xa_2 - a_2^2 - y^2 + 2yb_2 - b_2^2 = c_1 - c_2$$

$$-2xa_1 + 2xa_2 + a_1^2 + b_1^2 - 2yb_1 + 2yb_2 - a_2^2 - b_2^2 = c_1 - c_2$$

$$2x(a_2 - a_1) + 2y(b_2 - b_1) + a_1^2 + b_1^2 - a_2^2 - b_2^2 = c_1 - c_2$$

Further taking constants to one side;

$$2x(a_2 - a_1) + 2y(b_2 - b_1) = c_1 - c_2 - a_1^2 - b_1^2 + a_2^2 + b_2^2 \\ = C$$

Where;

$$C = c_1 - c_2 - a_1^2 - b_1^2 + a_2^2 + b_2^2$$

Therefore, the line that appears at the end is  $\boxed{2x(a_2 - a_1) + 2y(b_2 - b_1) = C}$

## 6. a

A complex plane is a geometric representation of the complex numbers established by the real axis and the orthogonal imaginary axis.

[Comment](#)

### Step 2 of 3 ^

To show: That thinking of a plane as the complex plane, describe the set of constructible points as complex numbers.

Thinking of the plane as the complex plane means that a complex number;

$$z = x + iy$$

This is constructible if the coordinate point  $(x, y)$  is constructible in the plane.

Now, a real number  $x$  is constructible if the point  $(0, x)$  is constructible.

If  $z$  is constructible, then taking the line passing through  $(x, y)$  that is parallel to the vertical axis allows to construct  $(x, 0)$  and by taking the parallel to the  $x$ -axis passing through  $(x, y)$  gives  $(0, y)$ .

Conversely,

If  $x$  and  $y$  are constructible reals, then taking the perpendicular at  $(x, 0)$  to the  $x$ -axis, the perpendicular at  $(0, y)$  to the  $y$ -axis, their intersection point is constructible.

Hence,  $z$  is constructible.

**Thus,  $z = x + iy$  is constructible if and only if both  $x$  and  $y$  are constructible.**

## Section 6

### 1. a

Consider  $F$  be a field of characteristic 0, and let  $f'$  be the derivative of  $f \in F[x]$  also, let  $g$  be an irreducible polynomial that is a common divisor of  $f$  and  $f'$ .

To prove:  $g^2$  divides  $f$

For the proof, consider:

$$f = gx$$

For some  $x \in F[x]$

Now, from the result which states that;

The derivative of a polynomial  $f$  with coefficients in a field  $F$  defined by the calculus formula;

$$(a_n x^n + \dots + a_1 x + a_0)' = na_n x^{n-1} + \dots + 1a_1$$

The integer coefficients are interpreted in  $F$  using the unique homomorphism  $\mathbb{Z} \rightarrow F$  then the product rule will be;

$$(fg)' = fg' + f'g$$

And, the chain rule is;

$$(f \circ g)' = (f' \circ g)g'$$

Then,

$$\begin{aligned} f' &= (gx)' \\ &= gx' + g'x \end{aligned}$$

Now, if;

$$g \mid f'$$

Then,

$$g \mid g'x$$

This is, prime since, it is irreducible

Thus, from irreducibility;  $g \mid g'$  or  $g \mid x$

But,  $g \nmid g'$  this is because;

$$\deg g' < \deg g$$

Hence,  $x = gy$  for some  $y \in F[x]$

Hence,

$$f = g^2 y$$

That is;

$$g^2 \mid f$$

2. a



Let  $F$  be a field of characteristic zero, let  $f'$  denote the derivative of a polynomial  $f$  in  $F[x]$  and let  $g$  be an irreducible polynomial that is a common divisor of  $f$  and  $f'$ .

To prove:  $g^2$  divides  $f$

**a.**

Let  $\alpha \neq 1$  satisfy the equation;

$$x^n - 1 = 0$$

Then  $\alpha$  is an  $n$ th root of unity and;

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] = n - 1$$

That is this is the degree of the corresponding cyclotomic polynomial

Suppose  $\alpha$  is in  $\mathbb{Q}(\sqrt{d})$

Then a chain can be created of extensions;

$$\mathbb{Q} : \mathbb{Q}(\alpha) : \mathbb{Q}(\sqrt{d})$$

Now, it is known that;

$$[\mathbb{Q}(\sqrt{d}) : \mathbb{Q}] = 2$$

And,

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] \neq 1$$

So, this implies that it must have;

$$\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\alpha)$$

Thus,  $n$  must be 3 so, that is look for only 3<sup>rd</sup> roots of unity.

That is;  $F(\sqrt{x})$  contains square roots of all elements of the form  $xp^2$

**Therefore,  $g^2$  divides  $f$**

To classify: the quadratic extensions of  $\mathbb{Q}$

**b.**

Now, from the above explanation it can be said that;

$$\mathbb{Q}(\sqrt{x}) = \mathbb{Q}(\sqrt{y})$$

If and only if;

$$xp^2 = yq^2; p, q \in \mathbb{Q}$$

**Thus, the distinct quadratic extensions of  $\mathbb{Q}$  are of the form of  $\mathbb{Q}(\sqrt{x})$  where  $x$  is not a square and has no square factor.**

3. a

An  $n$ th root of unity is defined as the case when  $n = 1, 2, 3, \dots$  is a number  $z$  satisfying the equation;

$$z^n = 1$$

[Comment](#)

Step 2 of 3 ^

To determine: All the quadratic number fields  $\mathbb{Q}(\sqrt{d})$  which contain a primitive  $p$ th root of unity, for some prime  $p \neq 2$

Let  $\zeta \neq 1$

This satisfies that;

$$x^p - 1 = 0$$

Now,

$$[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1$$

Since, this is a cyclotomic extension

If  $\zeta \in \mathbb{Q}(\sqrt{d})$ , this implies that a tower can be created as follows;

$$\begin{aligned} \mathbb{Q}(\sqrt{d}) &\rightarrow \mathbb{Q}(\zeta) \\ &\rightarrow \mathbb{Q} \end{aligned}$$

Now;

$$[\mathbb{Q}(\sqrt{d}) : \mathbb{Q}] = 2$$

And;

$$[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1$$

So, the only way this tower can occur is if;

$$\begin{aligned} \mathbb{Q}(\sqrt{d}) &= \mathbb{Q}(\zeta) \\ p &= 3 \end{aligned}$$

And since;

$$\mathbb{Q}(\sqrt{-3}) = \mathbb{Q}(\zeta_3)$$

But this is the only case that quadratic number field contains a  $p$ th root of unity for  $p > 2$

**Therefore, all the quadratic number fields  $\mathbb{Q}(\sqrt{d})$  which contain a primitive  $p$ th root of unity, for some prime  $p \neq 2$**

## Section 7

1. a

Group is defined as the algebraic structure consisting of a set of elements which is equipped with the operations that combines any two elements to form the third element.

[Comment](#)

Step 2 of 3 ^

To identify: The group  $\mathbb{F}_4$

Here,  $\mathbb{F}_4$  have characteristic 2, which is to say that every element  $x$  satisfies;

$$2x = 0$$

That is;

$$x + x = 0$$

So,  $\mathbb{F}_4^+$  is a group of order 4 in which each element is its own additive inverse, that is;

$$(\mathbb{Z}/2\mathbb{Z})^2$$

Also, the  $\mathbb{F}_4^+$  is isomorphic to;

$$\mathbb{F}_2[x]/(x^2 + x + 1)$$

Which consist of the elements  $0, 1, x, 1 + x$

Thus, the group of  $\mathbb{F}_4^+$  is  $(\mathbb{Z}/2\mathbb{Z})^2$

## 2. a

To determine: the irreducible polynomial of each of the elements of  $\mathbb{F}_8$  for;

$$\mathbb{F}_8 = \{0, 1, \beta, 1 + \beta, \beta^2, 1 + \beta^2, \beta + \beta^2, 1 + \beta + \beta^2\}$$

The polynomial;

$$P(X) = X^3 + X + 1$$

This is irreducible over  $\mathbb{F}_2$  else, it will have a factor of degree 1.

That is a root in  $\mathbb{F}_2$ , while;

$$\begin{aligned} P(0) &= P(1) \\ &= 1 \end{aligned}$$

Then,  $\mathbb{F}_8$  can be represented by all the triplets  $(a, b, c)$  of elements in  $\mathbb{F}_2$  or equally by all the polynomials of the form;

$$aX^2 + bX + c \in \mathbb{F}_2[X]$$

Then, addition and multiplication correspond to the addition and multiplication modulo  $P(X)$  of

$$(aX^2 + bX + c) \text{ in } \mathbb{F}_2[X]$$

Further consider the following;

$$\begin{aligned}(1,1,0) + (0,1,1) &= (X^2 + X) + (X + 1) \\ &= (1,0,1)\end{aligned}$$

And;

$$\begin{aligned}(1,1,0) \times (0,1,1) &= (X^2 + X)(X + 1) \bmod P(X) \\ &= (X^3 + X) \bmod P(X) \\ &= (0,0,1)\end{aligned}$$

Here,  $\beta$  is the residue class of  $X$ , which corresponds to the element  $(0,1,0)$

Since, by construction it can be said that;

$$P(\beta) = 0$$

It can be said that  $\mathbb{F}_8$ , is obtained from  $\mathbb{F}_2$  by adjoining a root of  $P$

Now, the elements in  $\mathbb{F}_8$  can be seen as quadratic polynomials in  $\beta$ , that is;

$$P(\beta) = 0$$

This means;

$$\beta^3 = \beta + 1$$

Now, to reduce the powers of  $\alpha$ , check that  $\alpha$  is a primitive element of  $\mathbb{F}_8$  also,  $\mathbb{F}_8^*$  contains;

$$\phi(1) = 1$$

This is an element of order 1

And;

$$\phi(7) = 6$$

This is elements of order 7, this implies that all elements different from 1 are generators of  $\mathbb{F}_8^*$ , the similar table has been shown below;

Powers of $\beta$	–	$\beta^0$	$\beta$	$\beta^2$	$\beta^3$	$\beta^4$	$\beta^5$	$\beta^6$
Polynomials in $\beta$	0	1	$\beta$	$\beta^2$	$\beta + 1$	$\beta^2 + \beta$	$1 + \beta + \beta^2$	$1 + \beta^2$

Therefore, the required irreducible polynomial is;

Powers of $\beta$	–	$\beta^0$	$\beta$	$\beta^2$	$\beta^3$	$\beta^4$	$\beta^5$	$\beta^6$
Polynomials in $\beta$	0	1	$\beta$	$\beta^2$	$\beta + 1$	$\beta^2 + \beta$	$1 + \beta + \beta^2$	$1 + \beta^2$

### 3. a

Fermat's little theorem states that if  $p$  is a prime number, then for integer  $a$ , the number  $a^p - a$  is an integer multiple of  $p$ , this expressed as;

$$a^p \equiv a \pmod{p}$$

[Comment](#)

Step 2 of 2 ^

To find: the 13<sup>th</sup> root of 2 in the field  $\mathbb{F}_{13}$

Now, from the above discussed Fermat's little theorem;

$$a^p \equiv a \pmod{p}$$

Here, for any prime  $p$  and any integer  $a$

So,  $\mathbb{Z}/13\mathbb{Z}$  this means;

$$2^{13} = 2$$

Thus,  $2^{13} = 2$  is the only 13<sup>th</sup> root of 2 in the field  $\mathbb{F}_{13}$

#### 4. a

To determine: the number of irreducible polynomials of degree 3 over  $\mathbb{F}_3$  and over  $\mathbb{F}_5$

Use the theorem which states that;

For prime  $p$  the irreducible factors of  $x^{p^3} - x$  are the monic irreducible polynomials in  $\mathbb{F}_p[x]$  with degree 1 or 3.

So,  $x, x-1, x-2, \dots, x-(p-1)$  are the linear irreducible polynomials

Hence, there are  $p(p-1)(p+1)/3$  monic irreducible polynomials of degree 3 in  $\mathbb{F}_p[x]$

Since, by theorem which states that;

Let  $K$  be a field of order  $q$ . The multiplicative group  $K^*$  of nonzero elements of  $K$  is a cyclic group of order  $q-1$

By the above result of the theorem there are  $p-1$  units in  $\mathbb{F}_p$

Also, there are  $p(p-1)^2(p+1)/3$  irreducible polynomials of degree 3 over  $\mathbb{F}_p$

That is irreducible polynomials for  $\mathbb{F}_3$ ;

$$\frac{3(3-1)^2(3+1)}{3} = \frac{3 \times 4 \times 4}{3} = 16$$

And, irreducible polynomials for  $\mathbb{F}_5$ ;

$$\frac{5(5-1)^2(5+1)}{3} = \frac{5 \times 16 \times 6}{3} = 160$$

Thus, for  $\mathbb{F}_3$  there are 16 irreducible polynomials and for  $\mathbb{F}_5$  there are 160 irreducible polynomial of degree 3

#### 5. a

Polynomial is defined as the expression of more than two algebraic terms, where the terms are sum of different variables of different powers.

[Comment](#)

Step 2 of 3 ^

To factor:  $x^9 - x$  and  $x^{27} - x$  in  $\mathbb{F}_3$

The monic polynomials of degree at most 3 over  $\mathbb{F}_3$  are;

$$\begin{aligned} &x, x+1, x-1, x^2+1, x^2+x-1, x^2-x-1, \\ &x^3-x+1, x^3-x-1, x^3+x^2-1, x^3-x^2+1, \\ &x^3+x^2+x+1, x^3+x^2+x-1, x^3+x^2-x+1, \\ &x^3-x^2+x+1, x^3-x^2+x-1, x^3-x^2-x-1 \end{aligned}$$

Since, the irreducible factors of a polynomial  $x^r - x$  over  $\mathbb{F}_3$  are the irreducible polynomials over  $\mathbb{F}_3$  whose degree divide  $r$ ;

$$\begin{aligned}
 x^9 - x &= x(x^8 - 1) \\
 &= x((x^4)^2 - (1)^2) \\
 &= x(x^4 - 1)(x^4 + 1) \\
 &= x((x^2)^2 - 1)(x^4 + 1) \\
 &= x(x^2 - 1)(x^2 + 1)(x^4 + 1) \\
 &= x(x+1)(x-1)(x^2 + 1)(x^4 + 1) \\
 &= x(x+1)(x-1)(x^2 + 1)(x^2 + x - 1)(x^2 - x - 1)
 \end{aligned}$$

And;

$$\begin{aligned}
 x^{27} - x &= x(x+1)(x-1)(x^3 + x^2 + x + 1)(x^3 + x^2 + x - 1) \dots \\
 &\dots (x^3 + x^2 - x + 1)(x^3 - x^2 + x + 1)(x^3 - x^2 + x - 1)(x^3 - x^2 - x - 1)
 \end{aligned}$$

6. a

Since,  $x^{24} - x$  will split as the linear, the quadratic and the degree-4 irreducibles over  $\mathbb{F}_2$

Over  $\mathbb{F}_{2^2}$ , the quadratic polynomial will split, while the degree-4 irreducibles will each split into two degree-2 ones.

Over, neither the degree-2 nor the degree-4 polynomials split at all so, the factorization is exactly over  $\mathbb{F}_2$

Now, computing;

$$\begin{aligned}
 x^{16} - x &= x^{16} + x \\
 &= x(x+1)(x^2 + x + 1)(x^4 + x + 1)(x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1)
 \end{aligned}$$

There is a unique degree-2 irreducible; therefore a degree-4 reducible polynomial is either;

$$(x^2 + x + 1)^2 = x^4 + x^2 + 1$$

Or, it has a root, having a root means either no constant term, or an even number of terms.

From above it can be said that  $\mathbb{F}_8$

Over  $\mathbb{F}_2$ , let  $\alpha$  be the root of;  $x^2 + x + 1$

And, that this polynomial must all 6 irreducibles of degree 2.

For checking whether the following have roots where;

$$\alpha^2 = \alpha + 1$$

And that;

$$\begin{aligned}
 f(\alpha) &= 0 \\
 f(\alpha + 1) &= 0
 \end{aligned}$$

Further computing;

$$\begin{aligned}
 x^{16} - x &= x^{16+x} \\
 &= x(x+1)(x+\alpha)(x+\alpha+1)(x^2 + \alpha x + 1)(x^2 + (\alpha+1)x + 1) \\
 &\quad (x^2 + x + \alpha)(x^2 + x + (\alpha+1))(x^2 + \alpha x + \alpha)(x^2 + (\alpha+1)x + (\alpha+1))
 \end{aligned}$$

7. a



Consider  $K$  be a finite field.

To prove; the product of the nonzero elements of  $K$  is  $-1$

For every nonzero  $a \in K$  there is  $a^{-1} \in K$  such that  $aa^{-1} = 1$  since  $K$  is a field.

In the product of nonzero elements of  $K$  that can be paired off each as  $aa^{-1}$  such that;

$$a \neq a^{-1}$$

So, that the product of all nonzero elements of  $K$  such that  $a^{-1} = a$

Claim: This is true if and only if  $a = \pm 1$

If  $a = a^{-1}$  then;

$$a^2 = 1$$

And, so  $a$  is a root of;

$$x^2 - 1 = (x+1)(x-1)$$

By using the proposition which states that;

The rings  $\mathbb{Z}[i]$  and the polynomial ring  $F[x]$  in one variable over a field  $F$  are unique factorisation domain.

Since,  $K[x]$  is unique factorization domain

It implies that  $a = \pm 1$

Thus, the product is equals to;

$$1 \times (-1) = (-1)$$

**Therefore, the product of the nonzero elements of  $K$  is  $-1$**

8. a

Consider the polynomials  $f(x) = x^3 + x + 1$  and  $g(x) = x^3 + x^2 + 1$  are irreducible over  $\mathbb{F}_2$ . Let  $K$  be the field extension obtained by adjoining a root of  $f$ , and let  $L$  be the extension obtained by adjoining a root of  $g$ .

To describe: An isomorphism from  $K$  to  $L$  and determine the number of such isomorphism

Let  $\alpha$  be a root of  $f$  and  $\beta$  one of  $g$ ;

Then;

$$\alpha^3 = \alpha + 1$$

And;

$$\beta^3 = \beta^2 + 1$$

Any field homomorphism;

$$\varphi: K \rightarrow L$$

This mapping is specified by  $\varphi(\alpha)$  satisfying that;

$$\varphi(\alpha^3) = \varphi(\alpha + 1)$$

So, by the mapping property of the quotient ring this states that;

Let  $f: R \rightarrow R'$  be a ring homomorphism with kernel  $K$  and let  $I$  be another ideal. Let;

$$\pi: R \rightarrow \bar{R}$$

Be the canonical map from  $R$  to  $\bar{R} = R/I$ . If  $I \subset K$ , there is a unique homomorphism;

$$\bar{f}: \bar{R} \rightarrow R' \text{ Such that } \bar{f}\pi = f$$

So, from the above result it can be said that;

$$\varphi(\alpha) = a + b\beta + c\beta^2$$

Where;

$$\begin{aligned} \varphi(\alpha^3) &= (a^2 + b^2\beta^2 + c^2\beta^4)(a + b\beta + c\beta^2) \\ &= (a^3 + ac^2 + b^3 + b^2c + bc^2) + (a^2b + ac^2 + b^2c + bc^2 + c^2)\beta + \\ &\quad (a^2c + ab^2 + ac^2 + b^3 + b^2c + c^3)\beta^2 \end{aligned}$$

This must be equal to  $\varphi(\alpha + 1)$

Claim:  $\varphi$  is an isomorphism if and only if ;

$$(a, b, c) \in \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$$

Then, by proposition which states that;

Let  $\varphi: F[x] \rightarrow R'$  be a homomorphism to an integral domain  $R'$  and let  $P$  be the kernel of  $\varphi$ .

Either  $P$  is a maximal ideal, or  $P = (0)$

Now, from the above result it remains to show that  $\varphi$  is surjective if only if this is true.

Suppose  $\varphi$  is surjective, then not both  $b, c$  are zero by the above.

Now, matching  $1, \beta, \beta^2$  terms, that is if;

$$b = 0$$

Then;

$$a = c$$

$$= 1$$

Now, if;

$$c = 0$$

Then;

$$a = b$$

$$= 1$$

And if;

$$b = c$$

$$= 1$$

Then,

$$a = 0$$

Conversely;

Since, in the first case;

$$\varphi(\alpha + 1) = \beta$$

In the second case;

$$\varphi(\alpha^2 + \alpha + 1) = \beta$$

And, in the third case;

$$\varphi(\alpha^2 + 1) = \beta$$

It can be seen that in each case  $\varphi$  is surjective by using that  $\varphi$  is a homomorphism.

9. a

a.

To determine: the number of monic irreducible polynomials of degree 2 in  $F[x]$

Suppose  $F$  is a field such that;

$$|F| = q$$

Take the degree as 2

Now;

$$|\mathbb{F}_{q^2} : \mathbb{F}_q| = p$$

Therefore, there could be no sub extensions and every irreducible polynomial that divides  $g$  must be of order 2 or 1.

Since, each linear polynomial over  $\mathbb{F}_q$  divides  $g$  and from the fact that  $g$  has distinct roots, then there are exactly  $q$  different linear polynomials that divide  $g$

So, denote the number of monic irreducible polynomials of degree 2 by  $x$ , then;

$$xp + q = q^2$$

That is;

$$\begin{aligned} x &= \frac{q^2 - q}{2} \\ &= \frac{q(q-1)}{2} \end{aligned}$$

Therefore, the number of monic irreducible polynomials of degree 2 in  $F[x]$  is;

$\frac{q(q-1)}{2}$
--------------------

---

b.

Let  $f(x)$  be an irreducible polynomial of degree 2 in  $F[x]$

To prove: that  $K = F[x]/(f)$  is a field of order  $p^2$ , and that its element has the form  $a + b\alpha$ , where  $a, b$  are in  $F$  and  $\alpha$  is a root of  $f$  in  $K$

Since,  $F$  is finite then it must have characteristic  $p$  for some prime  $p$

Here, consider;

$$p = 2$$

Where,  $p$  denotes a prime number

Thus, the prime subfield  $K$  of  $F$  is isomorphic to  $\mathbb{F}_p$

Then,  $F$  has  $p^2$  elements, where  $p$  is the characteristic of  $F$

**Therefore,  $K = F[x]/(f)$  is a field of order  $p^2$**

Now, to show that the elements of  $K = F[x]/(f)$  have the form  $a + b\alpha$

Now, let  $\phi(a + b\alpha) = a + b\alpha$

Then;

$$a^2 - b^2\alpha = a + b\alpha$$

That is;

$$b^2 = -b$$

This means that  $a$  can take any value in  $\mathbb{F}_2, b = 0$

Thus,  $\mathbb{F}$  is a fixed field of  $\phi$

**Therefore, the elements of  $K = F[x]/(f)$  have the form  $a + b\alpha$  and thus, for  $b \neq 0$  every element is the root of an irreducible quadratic polynomial in  $F[x]$**

c.

To show: that every polynomial of degree 2 in  $F[x]$  has a root in  $K$

Over  $\mathbb{F}_2$ , let  $\alpha$  be the root of  $x^2 + x + 1$

And, this must get all the six irreducibles of degree 2.

Now, it can be seen clearly that;

$$\alpha^2 = \alpha + 1$$

And;

$$f(\alpha) = 0$$

If and only if;

$$f(\alpha + 1) = 0 \text{ in } K$$

**This proves that every polynomial of degree 2 in  $F[x]$  has a root in  $K$**

d.

To show: that all the fields  $K$  constructed as above for a given prime  $p$  are isomorphic

Since, every degree 2 extension is normal.

And, if;

$$[K : F] = 2$$

Then;

$$K = F(\alpha)$$

Where,  $\alpha$  is a root of an irreducible polynomial  $f$  over  $K$ .

Thus,  $f$  splits in  $K$ , so  $K/F$  is normal

**Therefore, all the fields  $K$  constructed as above for a given prime  $p$  are isomorphic**

10. a

Consider  $F$  be a field, and let  $f(x)$  be a non-constant polynomial whose derivative is the zero polynomial.

To prove: That  $f$  cannot be irreducible over  $F$

For proving the required result first proves the lemma which states that;

In a finite field of order  $p^r$ , every element is a  $p$ th power.

The proof of the lemma is as follows;

Let  $a \in F$ ;

If  $a = 0$

Then the proof is done.

So, consider;

$$|F^\times| = p^r - 1$$

Now, by theorem which states that;

Let  $p$  be a prime integer and let  $q = p^r$  be a positive power of  $p$ . Let  $K$  be a field of order  $q$ .

The multiplicative group  $K^\times$  of nonzero elements of  $K$  is a cyclic group of order  $q - 1$

Hence,

$$a^{p^r-1} = 1$$

And;

$$a^{p^r} = a$$

Thus,  $b = a^{p^{r-1}}$  satisfies  $b^p = a$

Considering the proof of the lemma, the proof of the required result follows as given below:

$$\text{Let } f(x) = \sum_{i=0}^n a_i x^i$$

Then, if;

$$f'(x) = \sum_{i=1}^n i a_i x^{i-1} \\ = 0$$

Then;

$$i a_i = 0 \forall i \geq 1$$

Since,  $F$  is a domain, this implies that either;

$$a_i = 0 \text{ or } i \equiv 0 \pmod{p}$$

Thus;

$$f(x) = \sum_{j=0}^m a_j x^{p^j}, \text{ for some } a_j \in F$$

Here;

$$p = \text{char} F$$

Now, by using the above proved lemma there exist  $b_j$  such that  $b_j^p = a_j$

Thus;

$$f(x) = g(x)^p$$

Where;

$$g(x) = \sum_{j=0}^m b_j x^j$$

Therefore,  $f(x)$  is not irreducible.

**Hence,  $f$  cannot be irreducible over  $F$**

11. a

Consider  $f = ax^2 + bx + c$  with  $a, b, c$  in a ring  $R$

To show: The polynomial ring  $R[x]$  that is generated by  $f$  and  $f'$  contains the discriminant, the constant polynomial  $b^2 - 4ac$

For the proof first differentiate the function as given below:

$$f' = 2ax + b$$

By using the Euclidean algorithm divide  $f$  by  $f'$ , that is shown below:

$$\begin{array}{r} \frac{1}{2}x + \frac{b}{4a} \\ 2ax + b \overline{) ax^2 + bx + c} \\ \underline{ax^2 + \frac{1}{2}bx} \phantom{+ c} \\ \frac{1}{2}bx + c \\ \underline{\frac{1}{2}bx + \frac{b^2}{4a}} \\ -\frac{b^2}{4a} + c \end{array}$$

Now, by applying the division algorithm;

$$\begin{aligned} ax^2 + bx + c &= \left(\frac{1}{2}x + \frac{b}{4a}\right)(2ax + b) + \left(\frac{-b^2 + 4ac}{4a}\right) \\ &= \left(\frac{2ax + b}{4a}\right)(2ax + b) - \left(\frac{b^2 - 4ac}{4a}\right) \\ 4a^2x^2 + 4abx + 4ac &= (2ax + b)(2ax + b) - (b^2 - 4ac) \\ 4af &= f'^2 - (b^2 - 4ac) \end{aligned}$$

That is:

$$\begin{aligned} b^2 - 4ac &= -4af + f'^2 \\ &\in (f, f') \\ &\subset R[x] \end{aligned}$$

Therefore, the polynomial ring  $R[x]$  that is generated by  $f$  and  $f'$  contains the discriminant, the constant polynomial  $b^2 - 4ac$

## 12. a

A prime integer is defined as the integer that can be divisible by 1 or itself and also, the integer must be greater than 1.

[Comment](#)

Step 2 of 3 ^

Let  $p$  be a prime integer and let  $q = p^r$  and  $q' = p^k$

To find: For what values of  $r$  and  $k$  does  $x^{q'} - x$  divide  $x^q - x$  in  $\mathbb{Z}[x]$

First suppose  $f$  divides  $x^{q'} - x$

Let  $\alpha$  be a root of  $f$  in the splitting field of  $f$  over  $\mathbb{F}_q$

Then;  $\alpha^{q'} = \alpha$  so,  $\alpha \in \mathbb{F}_{q'}$

Thus,  $\mathbb{F}_q(\alpha)$  is a sub field of  $\mathbb{F}_{q'}$

Since;

$$[F_q(\alpha) : \mathbb{F}_q] = m$$

And;

$$[F_{q'} : \mathbb{F}_q] = k$$

Then;

$$[F_{q'} : \mathbb{F}_q(\alpha)] = k$$

So,  $r$  divides  $k$



Conversely suppose  $r$  divides  $n$

Then,  $\mathbb{F}_{q^r}$  contains  $\mathbb{F}_{q^k}$  as a subfield

Let  $\alpha$  be a root of  $f$  in the splitting field of  $f$  over  $\mathbb{F}_q$

Then;

$$[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = r$$

And, so;

$$\mathbb{F}_q(\alpha) = \mathbb{F}_{q^r}$$

Thus;

$$\alpha \in \mathbb{F}_{q^r}$$

Hence;

$$\alpha^{q^r} = \alpha$$

And, so  $\alpha$  is a root of  $x^{q^r} - x \in \mathbb{F}_q[x]$

This implies that  $f$  divides  $x^{q^r} - x$

**Thus, for all values of  $r$  and  $k$   $x^{q^r} - x$  divide  $x^q - x$  in  $\mathbb{Z}[x]$**

13. a

A group  $G$  called cyclic if there exists an element  $x$  in  $G$  such that:

$$\begin{aligned} G &= \langle x \rangle \\ &= \{x^n \mid n \text{ is an integer}\} \end{aligned}$$

[Comment](#)

Step 2 of 3 ^

Consider a finite cyclic subgroup

$$H \leq F^\times$$

This is a finite abelian group

Now, by using the theorem which states that;

A finitely generated abelian group  $V$  is a direct sum of cyclic subgroups  $C_{d_1} \oplus \dots \oplus C_{d_k} \oplus L$  where the order  $d_i$  of  $C_{d_i}$  is greater than 1 and  $d_i$  divides  $d_{i+1}$  for  $i = 1, \dots, k-1$

Thus,

$$H \approx A_1 \oplus \dots \oplus A_n$$

Where,  $A_i$  are cyclic groups and  $|A_i|$  divides  $|A_{i+1}|$

$$\text{Now, let } a = |A_n|$$

Then;

$$x^a = 1 \text{ for every } x \in H$$

This means that every  $x \in H$  is a root of  $x^a - 1$

Now,  $x^a - 1$  has at most  $a$  roots so,  $a \geq |H|$

On the other hand  $a \leq |H|$  by the decomposition;

$$a = |H| \text{ and } n = 1$$

Thus,  $H$  is cyclic

14. a

The Euler function  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  is a mapping associating to each positive integer  $n$  then the number  $\phi(n)$  of integers  $m$  is relatively prime to  $n$

[Comment](#)

Step 2 of 3 ^

To find: A formula in terms of the Euler  $\phi$  function for the number of irreducible polynomials of degree  $n$  over the field  $\mathbb{F}_p$

Define the Euler function  $\phi$  as;

$$\phi(n) = \{t\}; 1 \leq t \leq n \text{ and } \gcd(t, n) = 1$$

That is;

$$\begin{aligned}\phi(2) &= \{1\} \\ &= 1\end{aligned}$$

And;

$$\begin{aligned}\phi(3) &= \{1, 2\} \\ &= 2\end{aligned}$$

And;

$$\begin{aligned}\phi(4) &= \{1, 3\} \\ &= 2\end{aligned}$$

And so on

Now, consider the cyclic group  $\mathbb{F}_p^*$

This cyclic group  $\mathbb{F}_p^*$  contains  $\phi(p-1)$  primitive elements, where  $\phi$  is Euler's function, the number of integers less than and relatively prime to  $p-1$

If the integer  $n$  has the prime factorization  $p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$  then;

$$\phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_r}\right)$$

Hence, the formula in terms of the Euler  $\phi$  function for the number of irreducible

polynomials of degree  $n$  over the field  $\mathbb{F}_p$  is  $\boxed{\phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_r}\right)}$

## Section 8

1. a

To prove: That every finite extension of a finite field has a primitive element

Assume  $q > 2$

Let;

$$h = p_1^{e_1} p_2^{e_2} \cdots p_m^{e_m}$$

This is the prime factor decomposition of;

$$h = q - 1$$

For every  $i$ , the polynomial  $x^{h/p_i} - 1$  has at most  $h/p_i$  roots in  $GF(q)$

Hence, there is at least one nonzero element in  $GF(q)$  that is not a root of this polynomial.

Let  $a_i$  be such an element and set;

$$b_i = a_i^{h/(p_i^{e_i})}$$

That is;

$$b_i^{p_i^{e_i}} = 1$$

And, the order of  $b_i$  is a divisor of  $p_i^{e_i}$

Conversely;

$$b_i^{p_i^{e_i-1}} = a_i^{h/(p_i^{e_i})} \neq 1$$

And so the order of  $b_i$  is  $p_i^{e_i}$

Claim: That the element;

$$b = b_1 b_2 \dots b_m$$

This has order  $h$

Suppose that the order of  $b$  is a proper divisor of  $h$  and is therefore a divisor at least one of the  $m$  integers;

$$h/p_i, 1 \leq i \leq m$$

Say,  $h/p_1$

Then;

$$1 = b^{h/p_1} = b_1^{h/p_1} b_2^{h/p_1} \dots b_m^{h/p_1}$$

Now, for;

$$1 < i$$

The  $p_i^{e_i}$  divides  $h/p_1$  and, hence;

$$b_i^{h/p_1} = 1$$

Therefore,

$$b_1^{h/p_1} = 1$$

This implies that the order of  $b_1$  must divide  $h/p_1$

This is a contradiction

**Thus, every finite field has a primitive element.**

2. a

To determine: All primitive elements for the extension  $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$  of  $\mathbb{Q}$

It is easier to find the nonprimitive elements.

If  $\alpha$  is nonprimitive and nonrational, then;

$$\begin{aligned}\mathbb{Q} &\subset \mathbb{Q}[\alpha] \\ &\subset \mathbb{Q}(\sqrt{2}, \sqrt{3})\end{aligned}$$

By the assumption about  $\alpha$ , those fields are distinct. Since;

$$\begin{aligned}4 &= [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] \\ &= [\mathbb{Q}[\alpha] : \mathbb{Q}] [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}[\alpha]]\end{aligned}$$

Then each term in the product equals 2, and in particular,

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}[\alpha]] = 2$$

It is known that;

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$$

And that  $\sqrt{2} + \sqrt{3}$  satisfies the rational irreducible polynomial;

$$f(x) = (x - (\sqrt{2} + \sqrt{3}))(x + (\sqrt{2} + \sqrt{3}))(x - (-\sqrt{2} + \sqrt{3}))(x - (\sqrt{2} - \sqrt{3}))$$

Now, since,

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}[\alpha]] = 2, \sqrt{2} + \sqrt{3}$$

This satisfies a degree 2 polynomial;

$$g(x) \in \mathbb{Q}[\alpha]$$

This divides all polynomials where;

$$h(x) \in \mathbb{Q}[\alpha][x]$$

This have  $\sqrt{2} + \sqrt{3}$  as a root

In particular  $g(x)$  divides  $f(x)$  which means that  $g(x)$  is the product of two of the factors that make up  $f(x)$

First case:

If the two factors are;

$$(x - (\sqrt{2} + \sqrt{3}))(x + (\sqrt{2} + \sqrt{3}))$$

Then, multiplying that out, then the following is obtained;

$$\begin{aligned}g(x) &= x^2 - (2 + 3 + 2\sqrt{6}) \\ &\in \mathbb{Q}[\alpha][x]\end{aligned}$$

Hence,

$$\mathbb{Q}[\sqrt{6}] \subset \mathbb{Q}[\alpha]$$

And conclude that;

$$\mathbb{Q}[\sqrt{6}] = \mathbb{Q}[\alpha]$$

Since, each field has dimensions 2 over  $\mathbb{Q}$

This gives the nonprimitive elements  $a + b\sqrt{6}; a, b \in \mathbb{Q}$

Second case:

If the two factors are;

$$(x - (\sqrt{2} + \sqrt{3}))(x - (-\sqrt{2} + \sqrt{3}))$$

Then, multiplying that out, then the following is obtained;

$$g(x) = x^2 - 2\sqrt{3}x + 1 \\ \in \mathbb{Q}[\alpha][x]$$

And conclude that;

$$\mathbb{Q}[\sqrt{3}] = \mathbb{Q}[\alpha]$$

This gives the nonprimitive elements  $a + b\sqrt{3}; a, b \in \mathbb{Q}$

Third case:

If the two factors are;

$$(x - (\sqrt{2} + \sqrt{3}))(x - (\sqrt{2} - \sqrt{3}))$$

Then, multiplying that out, then the following is obtained;

$$g(x) = x^2 - 2\sqrt{2}x - 1 \\ \in \mathbb{Q}[\alpha][x]$$

And conclude that;

$$\mathbb{Q}[\sqrt{2}] = \mathbb{Q}[\alpha]$$

This gives the nonprimitive elements  $a + b\sqrt{2}; a, b \in \mathbb{Q}$

## Section 9

1. a

Let  $f(x)$  be a polynomial with coefficients in a field  $F$ .

To prove: That if there is a rational function  $r(x)$  such that  $r^2 = f$ , then  $r$  is a polynomial

Suppose  $f \neq 0$  for otherwise  $r$  is increasingly 0.

Let  $r = p(x)/q(x)$  for the co primes  $p$  and  $q$ ;

Then,

$$p^2 = fq^2$$

Now using the result that  $F[x]$  is a unique factorization domain

Since,

$$p \nmid q$$

Then

$$p^2 \mid f$$

And, so;

$$f = p^2 s \text{ For some } s \in F[x]$$

But then;

$$\begin{aligned} f q^2 &= p^2 s q^2 \\ &= p^2 \end{aligned}$$

And so;

$$s q^2 = 1 \text{ and } s, q \in F$$

And so,  $r \in F[x]$

Since, the units in  $F[x]$  are the constant polynomials, then for  $s, q \in F$

And hence,  $r \in F[x]$

Therefore, **if there is a rational function  $r(x)$  such that  $r^2 = f$ , then  $r$  is a polynomial**

## 2. a

Branch points is defined as the different values of  $t$  when  $\frac{\partial f(t, x)}{\partial x} = 0$

And, the gluing data is the rule described by the permutation  $\sigma_v$  of the indices  $1, \dots, n$  that sends  $i \rightarrow j$

To determine: the branch points and the gluing data for the Riemann surfaces of the polynomials

**a.**

Consider the polynomial;

$$x^2 - t^2 + 1$$

Let,

$$f(t, x) = x^2 - t^2 + 1$$

Now, differentiating with respect to  $x$ ;

$$\frac{\partial f}{\partial x} = 2x$$

Here,  $X$  is a two sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$2x = 0$$

$$x = 0$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, 0) = 0$$

$$-t^2 + 1 = 0$$

$$t^2 = 1$$

$$t = \pm 1$$

Therefore, the branch points are  $\boxed{t = 1, -1}$



Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$  ;

$$\frac{\partial f}{\partial t} = -2t \\ \neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a two-cycle

Since, there are two sheets; each of the permutation is the transposition  $(12)$

Hence, the gluing data that is the required permutation in this case is  $(12)$

**b.**

Consider the polynomial;

$$x^4 - t - 1$$

Let,

$$f(t, x) = x^4 - t - 1$$

Now, differentiating with respect to  $x$  ;

$$\frac{\partial f}{\partial x} = 4x^3$$

Here,  $X$  is a four sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$  ;

$$4x^3 = 0 \\ x = 0$$

Substitute this value in the function, that is  $f(t, x) = 0$  ;

$$f(t, 0) = 0 \\ -t - 1 = 0 \\ t = -1$$

Therefore, the branch points are  $t = -1$

Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$  ;

$$\frac{\partial f}{\partial t} = -1 \\ \neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a four-cycle

Since, there are four sheets; each of the permutation is the transposition  $(1234)$

Hence, the gluing data that is the required permutation in this case is  $(1234)$

**c.**

Consider the polynomial;

$$x^3 - 3tx - 4t$$

Let,

$$f(t, x) = x^3 - 3tx - 4t$$

Now, differentiating with respect to  $x$  ;

$$\frac{\partial f}{\partial x} = 3x^2 - 3t$$

Here,  $X$  is a three sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$3x^2 - 3t = 0$$

$$x^2 = t$$

$$x = \pm\sqrt{t}$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, \sqrt{t}) = 0$$

$$(\sqrt{t})^3 - 3t\sqrt{t} - 4t = 0$$

$$t\sqrt{t} - 3t\sqrt{t} - 4t = 0$$

$$-2t\sqrt{t} - 4t = 0$$

$$-2t(\sqrt{t} - 2) = 0$$

$$-2t = 0, \sqrt{t} - 2 = 0$$

$$t = 0, \sqrt{t} = 2$$

$$t = 0, t = 4$$

And;

$$f(t, -\sqrt{t}) = 0$$

$$(-\sqrt{t})^3 - 3t(-\sqrt{t}) - 4t = 0$$

$$-t\sqrt{t} + 3t\sqrt{t} - 4t = 0$$

$$2t\sqrt{t} - 4t = 0$$

$$2t(\sqrt{t} - 2) = 0$$

$$2t = 0, \sqrt{t} - 2 = 0$$

$$t = 0, \sqrt{t} = 2$$

$$t = 0, t = 4$$

Therefore, the branch points are  $t = 0, 4$

Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$ ;

$$\frac{\partial f}{\partial t} = -3x - 4$$

$$\neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a three-cycle

Since, there are three sheets; each of the permutation is the transposition  $(123)$

Hence, the gluing data that is the required permutation in this case is  $(123)$

d.

Consider the polynomial;

$$x^3 - 3x^2 - t$$

Let,

$$f(t, x) = x^3 - 3x^2 - t$$

Now, differentiating with respect to  $x$ ;

$$\frac{\partial f}{\partial x} = 3x^2 - 3x$$

Here,  $X$  is a three sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$3x^2 - 3x = 0$$

$$3x(x - 1) = 0$$

$$3x = 0, x - 1 = 0$$

$$x = 0, x = 1$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, 0) = 0$$

$$-t = 0$$

$$t = 0$$

And;

$$f(t, 1) = 0$$

$$1 - 3 - t = 0$$

$$-2 - t = 0$$

$$t = -2$$

Therefore, the branch points are  $t = 0, -2$

Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$ ;

$$\frac{\partial f}{\partial t} = -1$$

$$\neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a three-cycle

Since, there are three sheets; each of the permutation is the transposition  $(123)$

Hence, the gluing data that is the required permutation in this case is  $(123)$

e.

Consider the polynomial;

$$x^3 - t(t-1) = x^3 - t^2 + t$$

Let,

$$f(t, x) = x^3 - t^2 + t$$

Now, differentiating with respect to  $x$ ;

$$\frac{\partial f}{\partial x} = 3x^2$$

Here,  $X$  is a three sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$3x^2 = 0$$

$$x = 0$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, 0) = 0$$

$$-t(t-1) = 0$$

$$-t = 0, t-1 = 0$$

$$t = 0, t = 1$$

Therefore, the branch points are  $t = 0, 1$

Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$ ;

$$\frac{\partial f}{\partial t} = -2t + 1$$

$$\neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a three-cycle

Since, there are three sheets; each of the permutation is the transposition  $(123)$

Hence, the gluing data that is the required permutation in this case is  $(123)$

f.

Consider the polynomial;

$$x^3 - 3tx^2 + t$$

Let,

$$f(t, x) = x^3 - 3tx^2 + t$$

Now, differentiating with respect to  $x$ ;

$$\frac{\partial f}{\partial x} = 3x^2 - 6tx$$

Here,  $X$  is a three sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$3x^2 - 6tx = 0$$

$$3x(2x - 6t) = 0$$

$$3x = 0, 2x - 6t = 0$$

$$x = 0, x = 3t$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, 0) = 0$$

$$t = 0$$

And;

$$f(t, 3t) = 0$$

$$(3t)^3 - 3t(3t)^2 + t = 0$$

$$27t^3 - 9t^3 + t = 0$$

$$18t^3 + t = 0$$

That is;

$$t(18t^2 + 1) = 0$$

$$t = 0, t^2 = -\frac{1}{18}$$

So,  $t = 0$

Therefore, the branch points are  $t = 0$

Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$ ;

$$\frac{\partial f}{\partial t} = -3x^2 + 1$$
$$\neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a three-cycle

Since, there are three sheets; each of the permutation is the transposition  $(123)$

Hence, the gluing data that is the required permutation in this case is  $(123)$

[Comment](#)

Step 15 of 19 ^

g.

Consider the polynomial;

$$x^4 + 4x + t$$

Let,

$$f(t, x) = x^4 + 4x + t$$

Now, differentiating with respect to  $x$ ;

$$\frac{\partial f}{\partial x} = 4x^3 + 4$$

Here,  $X$  is a four sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$4x^3 + 4 = 0$$

$$x = -1$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, -1) = 0$$

$$1 + 4 + t = 0$$

$$t = -5$$

Therefore, the branch points are  $t = -5$

Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$ ;

$$\frac{\partial f}{\partial t} = 1$$

$$\neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a four-cycle

Since, there are four sheets; each of the permutation is the transposition (1234)

Hence, the gluing data that is the required permutation in this case is  $\boxed{(1234)}$

h.

Consider the polynomial;

$$x^3 - 3tx - t - t^2$$

Let,

$$f(t, x) = x^3 - 3tx - t - t^2$$

Now, differentiating with respect to  $x$ ;

$$\frac{\partial f}{\partial x} = 3x^2 - 3t$$

Here,  $X$  is a three sheeted covering of  $T$

Now, put,  $\frac{\partial f}{\partial x} = 0$ ;

$$3x^2 - 3t = 0$$

$$x = \pm\sqrt{t}$$

Substitute this value in the function, that is  $f(t, x) = 0$ ;

$$f(t, \sqrt{t}) = 0$$

$$(\sqrt{t})^3 - 3t\sqrt{t} - t - t^2 = 0$$

$$t\sqrt{t} - 3t\sqrt{t} - t - t^2 = 0$$

$$-2t\sqrt{t} - t - t^2 = 0$$

$$-t(2\sqrt{t} + 1 + t) = 0$$

$$-t = 0, 2\sqrt{t} + 1 + t = 0$$

$$t = 0, (\sqrt{t} + 1)^2 = 0$$

$$t = 0, \text{ Not possible}$$

And;

$$f(t, -\sqrt{t}) = 0$$

$$(-\sqrt{t})^3 - 3t(-\sqrt{t}) - t - t^2 = 0$$

$$-t\sqrt{t} + 3t\sqrt{t} - t - t^2 = 0$$

$$2t\sqrt{t} - t - t^2 = 0$$

$$-t(-2\sqrt{t} + 1 + t) = 0$$

$$-t = 0, 2\sqrt{t} - 1 - t = 0$$

$$t = 0, (\sqrt{t} - 1)^2 = 0$$

$$t = 0, t = 1$$

Therefore, the branch points are  $\boxed{t = 0, 1}$



Now, finding the gluing data that means the permutation;

For, this first differentiate the function with respect to  $t$ ;

$$\frac{\partial f}{\partial t} = -1 - 2t$$

$$\neq 0$$

So, the branch point exists on all the points

That is, the permutation of the sheets at each of these points contains a three-cycle

Since, there are three sheets; each of the permutation is the transposition  $(123)$

Hence, the gluing data that is the required permutation in this case is  $\boxed{(123)}$

### 3. a

Isomorphism class defines all the elements in a given set and which are isomorphic to each other.

[Comment](#)

Step 2 of 4 ^

a.

To determine: the number of isomorphism classes of function fields  $K$  of degree 3 over  $F = \mathbb{C}(t)$  that the ramified only at the points 1 and -1

For the proof consider  $(C, f)$  be a branched cover

Now, given a neighborhood  $U$  of a point;

$$P \in C$$

Then if  $P$  is greater than 1 then it is said to be a ramified point

It can be seen that on  $U$  the map  $f$  is one to one.

Now, by the compactness of  $C$ , there is a finite open cover of  $C$  by which  $f$

is one to one at all but finitely many points, and hence can be ramified finitely at 1 and -1

**Therefore, the number of isomorphism classes of function fields  $K$  of degree 3 over  $F = \mathbb{C}(t)$  that the ramified only at the points 1 and -1**

b.

To describe: the gluing data for the Riemann surface corresponding to each isomorphism class of fields as a pair of permutations

Here, determining the permutation is defined by the gluing data of a Riemann surface.

For this at each branch point  $p$  determine the permutation  $\sigma$  of the sheets that occurs when one circles that points.

Now, choose a base point  $b$  in the cut plane  $T$  and compute the  $n$  distinct roots of the polynomial  $f(b, x)$  numerically. Consider the root be  $\gamma_1, \dots, \gamma_n$  and the sheets be  $S_i$  that contains the root  $\gamma_i$

Also, consider a point  $b_v$  in the vicinity of a branch point  $p_v$ , then  $\gamma_i$  varies continuously.

To determine, follow a counterclockwise loop around  $p_v$ , as the loop crosses the cut  $C_v$ , the roots will have been permuted by  $\sigma_v$  when the path returns to  $b_v$

Doing like the gluing data that is, the pair of permutation is obtained.

**Hence, obtaining the gluing data for the Riemann surface corresponding to each isomorphism class of fields as a pair of permutations**

c.

To find: a polynomial  $f(t, x)$  such that  $K = F[t]/(f)$

Let  $F$  be a finite field

Now, construct an irreducible quadratic polynomial  $f$  such that  $f$  is a polynomial.

Then  $F[t]/(f)$  is a two dimensional algebra over  $F$

Since,  $F[X]$  is an principal ideal domain, then any irreducible element is prime and any non-zero prime ideal is maximal

Hence,  $F[t]/(f)$  is a field of degree 2 over  $F$

Now, let;

$$S = \{t^2 + at + b : a, b \in F\}$$

This is the set of all monic quadratic polynomials in  $F[t]$

It can be seen that the quadratic polynomial is reducible over  $F$  if and only if it has a root in  $F$

Thus, the polynomial is of the form;

$$f(t, x) = \{t^2 + at + b : a, b \in F\}$$

#### 4. a

If  $\delta$  satisfies some irreducible quadratic equation in a field  $F$ , then  $F(\delta)$  is a quadratic extension field of  $F$

[Comment](#)

Step 2 of 4 ^

To show: that up to isomorphism, a quadratic extension of  $F$  is described by the finite set  $\{p_1, \dots, p_k\}$  of its true branch points

A degree 2 extension  $K$  of  $F = \mathbb{C}(t)$  can be obtained by adjoining the square root, say;

$$K = F[x]/(x^2 - d)$$

Where,  $d$  is a rational number in  $t$

Now, change  $d$  by a square factor  $r^2$ , because;

$$F[x]/(x^2 - d) \approx F[y]/(y^2 - r^2 d)$$

By the isomorphism;

$$rx \leftrightarrow y$$

Further, clear out the denominator of the rational function  $d$ , and also assume that  $d$  is a square free polynomial in  $t$ , that is;

$$d(t) = (t - a_1) \cdots (t - a_k)$$

This corresponds bijectively to finite sets of points  $p_1, \dots, p_k$  in the complex plane  $T$ ,  $p_i$  being the point  $t = a_i$

Consider, a two sheeted branched coverings  $X$  of  $T$  are also determined by gluing data at finite sets of points  $\{p_1, \dots, p_k\}$ , where  $p_i$  are the branch points.

Since, there are two sheets, the gluing data for  $p_i$  is one of the permutations (1) or (12)

If the permutation is trivial, ignore the point  $p_i$

Now, assume that the permutation at each point  $p_i$  is transposition (12)

In this way, isomorphism classes of two sheeted branched coverings also correspond bijectively to finite sets of points.

Claim: to show that if  $X$  is a Riemann surface of the field extension  $K$ ;

$$x^2 = (t - a_1) \cdots (t - a_k)$$

Then  $X$  is a point  $p_i : t = a_i$  are the true branch points of  $X$ , this means that the permutations of the sheets at each point is  $(12)$

Now, assume that this is true for the field extension;

$$K : z^2 = t$$

Let  $Z$  be the Riemann surface of the equation;

$$z^2 = t - a_1$$

And, let  $Y$  be the Riemann surface of the equation  $y^2 = (t - a_2) \cdots (t - a_k)$

By induction  $p_2, \dots, p_k$  are the true branch points of  $Y$ , and  $p_1$  is the branch point of  $Z$

The equation  $x = yz$  associates a point of  $X$  to a pair of points  $y \in Y$  and  $z \in Z$

Now, if  $y_0, z_0$  are the points of  $Y$  and  $Z$  respectively and follow  $t$  along the path, obtaining paths  $y(\theta)$  and  $z(\theta)$

Since,  $Y$  is not branched at  $p_1$ ;

$$\begin{aligned} y(2\pi) &= y(0) \\ &= y_0 \end{aligned}$$

While since,  $Z$  is branched here,  $z(2\pi) \neq z(0)$

So, let  $x(\theta) = y(\theta)z(\theta)$

Then;

$$x(2\pi) \neq x(0)$$

This means  $X$  is branched at  $p_1$

Therefore,  $X$  is branched at other points too.

**Hence, proved that up to isomorphism, a quadratic extension of  $F$  is described by the finite set  $\{p_1, \dots, p_k\}$  of its true branch points**

## 5. a

A branch point of any multi-variable function is defined as the point on the function which is discontinuous when going around an arbitrarily small circuit around this point.

[Comment](#)

Step 2 of 6 ^

To write: a computer program that determines the branch points  $p_i$  and the permutations  $\sigma_i$  for the Riemann surface of a given polynomial

For any given polynomial, first find the derivative with respect to  $x$  and put it equals to 0 then substitute the value of  $x$  in the polynomial to obtain the value of  $t$ . This value of  $t$  is defined as the branch points.

### Program Plan:

At first read coefficient of  $x^3, x^2$  and  $x$  from user and store it in array.

After that read coefficient of  $t$  and  $c$  from user.

After that use multiple if statements to check coefficient of  $t, c, x^3, x^2$  and  $x$ .

Use nested if statements to check whether coefficient is 1.

Use multiple if statements to determine maximum number of sheets.

After that call rootfind function to determine the roots

After finding the root at last find the branches point

## Section 10

1. a

**Solution:** We will propose to prove that the subset of  $\mathbb{C}$  consisting of the algebraic numbers is algebraically closed.

Let  $\overline{\mathbb{Q}}$  denote the algebraic numbers.

Let us start with some lemmas.

**Lemma-1:** The number  $a \in \mathbb{C}$  is algebraic if and only if the vector space  $V$  over  $\mathbb{Q}$  is defined by

$$V := \langle 1, a, a^2, a^3, \dots \rangle$$

is finite-dimensional.

**Proof:** Let us assume the dimension of  $V$  over  $\mathbb{Q}$  is  $d$ . Then the set  $\{1, a, a^2, \dots, a^d\}$  containing  $d + 1$  element is linearly dependent over  $\mathbb{Q}$ .

It follows that  $a$  satisfies an equation of degree less than equals  $d$ .

Thus  $a$  is algebraic over  $\mathbb{C}$ .

Conversely let us assume

$$a^n + a_1 a^{n-1} + \dots + a_{n-1} a + a_n = 0.$$

This follows that

$$a^n = -a_1 a^{n-1} - \dots - a_{n-1} a - a_n.$$

Therefore we have  $a^n \in \langle 1, a, a^2, a^3, \dots, a^{n-1} \rangle$  By similar manner we have

$$a^{n+1}, a^{n+2}, \dots \in \langle 1, a, a^2, a^3, \dots, a^{n-1} \rangle$$

This follows that

$$V = \langle 1, a, a^2, a^3, \dots, a^{n-1} \rangle$$

Hence  $V$  is finitely generated.

This completes the proof of first lemma.

**Lemma-2:** The number  $a \in \mathbb{C}$  is algebraic if and only if there exists a finite-dimensional vector space  $V$  of  $\mathbb{C}$  such that

$$aV \subset V.$$

**Proof:** By Lemma-1 we can take

$$V = \langle 1, a, a^2, \dots \rangle.$$

So we have the if one.

Now conversely let, dimension of  $V$  over  $\mathbb{Q}$  is  $d$ .

Let us choose  $v \in V$  such that  $v \neq 0$ . Then the  $d + 1$  elements of the set  $\{v, av, a^2v, \dots, a^dv\}$  are linearly dependent over  $\mathbb{Q}$ .

It follows that  $a$  satisfies an equation of degree less than equals  $d$ .

Thus  $a$  is algebraic over  $\mathbb{C}$ .

This completes the roof of second lemma.

Now we claim that  $\overline{\mathbb{Q}}$  is algebraically closed.

That is, if  $a \in \mathbb{C}$  satisfies an equation

$$x^n + c_1x^{n-1} + \dots + c_n = 0, \text{ where } c_i \in \overline{\mathbb{Q}}$$

then  $a \in \overline{\mathbb{Q}}$ .

For  $i = 1, 2, \dots, n$ , let us assume  $V_i$  be a finite-dimensional (but non-zero) vector space such that

$$c_i V_i \subset V_i.$$

And then consider

$$V_0 := \langle 1, a, a^2, \dots, a^{n-1} \rangle.$$

Let  $V$  be the vector space spanned by the products  $a^i v_1 \dots v_{n-1}$ , where  $v_i \in V_i$ .

Then clearly notice that

$$aV \subset V.$$

Now we will assert that

$$a^{i+1} v_1 v_2 \dots v_{n-1} \in V.$$

This is immediate unless  $i = n - 1$ .

Where we have

$$a^n v_1 v_2 \dots v_n = \sum_{0 \leq i < n} a^i v_1 \dots v_{i-1} (c_i v_i) v_{i+1} \dots v_{n-1}.$$

Now notice that  $c_i v_i \in V_i$ . Therefore note that

$$a^n v_1 v_2 \dots v_n \in V.$$

And it follows that

$$aV \subset V.$$

Since  $V$  is finite-dimensional, from **Lemma-2** it follows that  $a \in \overline{\mathbb{Q}}$ .

This completes the proof.

## Result

Considering  $\overline{\mathbb{Q}}$  denote the algebraic numbers, we have proved that  $\overline{\mathbb{Q}}$  is algebraically closed.



To construct: an algebraically closed field that contains the prime field  $\mathbb{F}_p$

For proving the required result it is enough to show that any field  $K$  is contained in an algebraically closed field  $L$ .

For the proof first claim that;

The monic irreducible polynomials of degree greater than or equals to 1 in  $K[x]$  can be well ordered.

This is a consequence of the well-ordering theorem, but in the case of  $\mathbb{F}_p$  or any countable field it is possible to well-order these polynomials without the axiom of choice since, there are only countably many of them. Then, letting  $(f_i)$  be a well-ordering of these polynomials, and letting;

$$L_0 = K$$

And,  $L_i$  be the extension of  $L_{i-1}$  such that  $f_i$  splits completely which exists by the proposition which states that;

Let  $F$  be a field and let  $f(x)$  be a monic polynomial in  $F[x]$  of positive degree. There exists a field extension  $K$  of  $F$  such that  $f(x)$  splits completely in  $K$

Claim:  $L = \bigcup_i L_i$  is an algebraically closed field containing  $K$

Here, from the above explanation it remains to prove that  $L = \bigcup_i L_i$  is algebraically closed, but this is true since if;

$$f(x) \in L[x]$$

It has a root  $\alpha$

Now by the result that;

A root of a polynomial with coefficients that are algebraic over a field  $F$  is also algebraic.

Let  $f \in A[x]$  be non-constant, and let  $\alpha$  be a root of  $f$ .

It remains to show that  $\alpha$  is algebraic over  $F$ .

Suppose not and let  $a_0, \dots, a_n$  be the coefficients of  $f$ . Then;

$$\begin{aligned} [F(\alpha):F] &\leq [F(a_0, \dots, a_n):F] \\ &\leq n \times \prod_{i=0}^n m_i \\ &< \infty \end{aligned}$$

By corollary;

Where  $m_i$  is the degree of  $a_i$  over  $F$ .

This, there exists a polynomial in  $F[x]$  of degree  $\leq n \times \prod_{i=0}^n m_i$  with  $\alpha$  as a root.

This shows  $\alpha$  is algebraic over  $K$

**Hence, lies in  $L$  by construction.**

3. a



A loop is defined same as a circle that is the end of a loop is always connected to its beginning part.

[Comment](#)

## Step 2 of 3 ^

To show: that what happens to the loop as  $r$  varies

For the proof consider the rule;

$$y = f(x)$$

This defines a function from the complex  $x$ -plane to the complex  $y$ -plane

Now, let  $C_r$  denote a circle of radius  $r$  about the origin in the complex  $x$ -plane, written as;

$$x = re^{i\theta}$$

With,  $0 \leq \theta \leq 2\pi$

Further, consider the function defined by the polynomial as;

$$\begin{aligned} y &= x^n \\ &= r^n e^{in\theta} \end{aligned}$$

Since,  $n\theta$  runs from 0 to  $2\pi n$

The point  $y$  winds  $n$  times around the circle of radius  $r^n$

So, the radius of the path is  $r^n$  and the length of the leash is  $\frac{1}{10} r^n$

Since,  $f$  is a continuous function, the image  $f(C_r)$  will vary continuously with  $r$

When the  $r$  is taken to be small then the function makes a small loop

As, the value of the  $r$  is taken large the loop also becomes bigger as compared to the taken value of the radius that is  $r$

But the for the different values of  $r$  being the value be small or big the loop never wind around the origin

This means at some point when the loop will be zero that will be defined as the root of that function, say that value be  $\alpha$ ;

$$f(\alpha) = 0$$

**Therefore, the range of the loop directly depends upon the value taken for the radius in the function, that is  $r$**

## 4. a

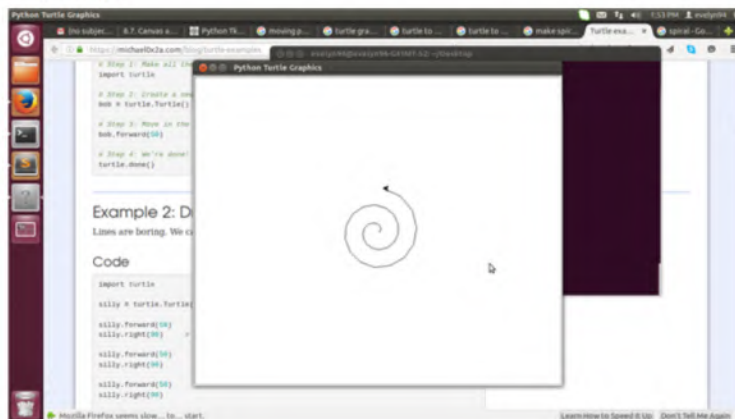
The function  $f$  is a continuous function, so the image  $f(C_r)$  will vary continuously with  $r$ . when the radius  $r$  is very small,  $f(C_r)$  makes a small loop around the constant term of  $f$ . The  $f(C_r)$  winds  $n$  times around the origin if  $r$  is large enough.

Now consider the following program that illustrates the variation of  $f(C_r)$  with radius  $r$ .

**Program:**

```
#Import the turtle library
from turtle import *
#Define the method spiral which takes three arguments
#Where n is the radius that varies accordingly
def spiral(n, angle, step):
#For loop is used to change the value continuously
for step in range(n):
#Move to the further steps
forward(step)
left(angle)
#call the function spiral
spiral(100,25,100)
x=input("Press any key to exit")
```

**Sample Output:**



## Miscellaneous Problem

1. a

A transcendental element is a complex number which is transcendental over the field of rational numbers, that is  $\mathbb{Q}$

[Comment](#)

Step 2 of 3 ^

Consider  $K = F[\alpha]$  is a field extension generated by a transcendental element  $\alpha$  and let  $\beta$  be an element of  $K$  that is not in  $F$ .

To prove:  $\alpha$  is algebraic over  $F(\beta)$

For the proof first suppose that;

$$\beta \in K \setminus F$$

Then;

$$\beta = p(\alpha)/q(\alpha)$$

Where,  $p(x), q(x) \in F[x]$

Thus,

$$\beta q(\alpha) - p(\alpha) = 0$$

So, from the above explanation  $\alpha$  is the root of  $\beta q(x) - p(x) \in F(\beta)[x]$

Hence, this implies that  $\alpha$  is algebraic over  $F(\beta)$

2. a

**Solution:** Let us assume  $f(x) = x^7 + x + 1$ .

We will factor  $f(x)$  in  $\mathbb{F}_7$ .

Now notice that by Fermat's little theorem

$$3^7 \equiv 3 \pmod{7}.$$

Therefore we have

$$f(3) \equiv 0 \pmod{7}.$$

Hence 3 is a root of  $f(x)$  in  $\mathbb{F}_7$ . Let us now consider an element  $\alpha$  in  $\mathbb{F}_7$  such that  $\alpha$  is a root of  $f(x)$  and  $\alpha \neq 3$ .

Let us now consider the Frobenius map  $\phi$  on  $\mathbb{F}_7$  as

$$\phi(x) = x^7, \quad \text{for } x \in \mathbb{F}_7.$$

Now if  $\alpha$  has minimal polynomial  $h$ , say, then  $\alpha^7$  must also be a root of  $h$ . But note that  $f(\alpha) = 0$ . Hence we have

$$\alpha^7 = -\alpha - 1.$$

Now look at the equation

$$(x - \alpha)(x + \alpha + 1) = x^2 + x - (\alpha^2 + \alpha).$$

Therefore our polynomial  $h(x)$  must be divided by  $x^2 + x - (\alpha^2 + \alpha)$ .

This follows that any irreducible polynomial of the form  $x^2 + x + t$  in  $\mathbb{F}_7[x]$  must be a factor, since it must have a root like  $\alpha$ .

Now we will find the irreducible polynomials of the form  $x^2 + x + t$  in  $\mathbb{F}_7[x]$ .

Now note that if the discriminant of  $x^2 + x + t$  is not in  $\mathbb{F}_7$ , then the polynomial  $x^2 + x + t$  will clearly be irreducible in  $\mathbb{F}_7[x]$ .

Now the discriminant of  $x^2 + x + t$  is given by

$$D = \sqrt{1 - 4t}.$$

Then if  $D = 0$  in  $\mathbb{F}_7$ , then  $x^2 + x + t$  is reducible.

So

$$D = 0 \implies t = 2, 0, 1, 5 \text{ in } \mathbb{F}_7.$$

Hence for  $t = 3, 4, 6$  in  $\mathbb{F}_7$ ,  $D \neq 0$ . Therefore for  $t = 3, 4, 6$  in  $\mathbb{F}_7$ , the polynomial  $x^2 + x + t$  is irreducible in  $\mathbb{F}_7[x]$ .

Hence for  $a, b \in \mathbb{F}_7$  we have

$$f(x) = (ax + b)(x^2 + x + 3)(x^2 + x + 4)(x^2 + x + 6).$$

It is obvious that  $a = 1$ , by comparing coefficients of  $x^7$ . Now by considering the constant term we have

$$72b \equiv 1 \pmod{7}$$

$$2b \equiv 1 \pmod{7}$$

$$8b \equiv 4 \pmod{7}$$

$$b \equiv 4 \pmod{7}.$$

Hence we have  $b = 4$ .

Thus the required factor of  $f(x)$  in  $\mathbb{F}_7$  is given by

$$x^7 + x + 1 = (x + 4)(x^2 + x + 3)(x^2 + x + 4)(x^2 + x + 6).$$

This completes the solution.

## Result

3 of 3

The required factor in  $\mathbb{F}_7$  is given by  $x^7 + x + 1 = (x + 4)(x^2 + x + 3)(x^2 + x + 4)(x^2 + x + 6)$ .

## 3. a

Irreducible polynomial is defined as the non-constant polynomial that cannot be factored into the product of two non-constant polynomials.

[Comment](#)

### Step 2 of 4 ^

Let  $f(x)$  be an irreducible polynomial of degree 6 over a field  $F$ , and that  $K$  be a quadratic extension of  $F$

To show: About the degrees of the irreducible factors of  $f$  in  $K[x]$

Let;

$$K = F[\alpha]$$

Where,  $\alpha$  is the root of some irreducible quadratic polynomial  $g \in F[x]$

Let  $\beta$  be a root of  $f$ , then;

$$[K : F] = 2$$

$$[F(\beta) : F] = 6$$

This divides  $[K(\beta) : F] \leq 12$

So, by using the corollary which states that;

Let  $\mathcal{K}$  be an extension field of  $F$ , let  $K$  and  $F'$  be subfields of  $\mathcal{K}$  that are finite extensions of  $F$ , and let  $K'$  denote the subfield of  $\mathcal{K}$  generated by the two fields  $K$  and  $F'$  together. Let;

$$[K' : F] = N$$

$$[K : F] = m$$

$$[F' : F] = n$$

Then  $m$  and  $n$  divide  $N$ , and  $N \leq mn$

So, from the above written result;

$$[K(\beta) : F] \in \{6, 12\}$$

In either case 3 divides the previous case,  $f$  splits into a product of polynomials of degree 3 over  $K$

[Comment](#)

#### Step 4 of 4 ^

In the previous case,  $f$  remains irreducible over  $K$ . These are the only possibilities for the irreducible factors of  $f$  in  $K$ , for the only possibility is that the factors of degree 3 factor further to include linear factors, which is impossible since  $\beta$  is of degree 6 over  $F$

## 4. a

Prime numbers are those numbers which are not divisible by any other numbers but is factor of 1 and itself.

[Comment](#)

#### Step 2 of 7 ^

Let  $p$  be an odd prime.

a.

To prove: That exactly half of the elements of  $F_p^*$  are squares and that if  $\alpha$  and  $\beta$  are non-squares, then  $\alpha\beta$  is a square

Let  $K$  be a field of order  $p^n$  for  $n$  odd, and consider the homomorphism;

$$\phi : K^* \rightarrow K^*$$

Such that;

$$\phi(x) = x^2$$

The kernel of this homomorphism is  $\{\pm 1\}$

That is the square root of unity

Thus, by the  $n^{\text{th}}$  isomorphism theorem;

The image is half the group.

**That is, half the elements are squared.**

b.

To prove: The same assertion for any finite field of odd order

Let  $F$  be the finite field of odd order and consider the group homomorphism;

$$\varphi: F^* \rightarrow F^*$$

Defined as;

$$\alpha \rightarrow \alpha^2$$

If  $\alpha \in \ker \varphi$  Then;

$$\alpha^2 = 1$$

Hence;

$$(x - \alpha) \mid x^2 - 1 \in F[x]$$

Since,  $F[x]$  is unique factorization domain, this is true by using the result which states that;

The rings  $\mathbb{Z}[i]$  and the polynomial ring  $F[x]$  in one variable over a field  $F$  are unique factorization domains.

This implies;

$$\alpha = \pm 1$$

Thus,  $|\ker \varphi| = 2$

Since,

$$\text{char } F \neq 2$$

This implies that;

$$1 \neq -1$$

And, so the set of squares  $\text{im } \varphi$  in  $F^*$  has order  $\frac{1}{2}|F^*|$  by the counting formula which states that;

Let  $\varphi: G \rightarrow G'$  be a homomorphism of finite groups. Then;

First is;

$$|G| = |\ker \varphi| \times |\text{im } \varphi|$$

Second is;

$$|\ker \varphi| \text{ divides } |G|$$

Third is;

$$|\text{im } \varphi| \text{ divides both } |G| \text{ and } |G'|$$

Now, let  $\beta$  be a non-square

Consider the map  $\mu: F^* \rightarrow F^*$  of sets defined by  $\alpha \rightarrow \alpha\beta$

This is a bijection and so to show  $\mu$  sends non-squares to squares, no it remains to show that  $\mu$  sends non-squares to squares

But this is true since if  $\alpha$  is a square and  $\alpha\beta$  is also a square, then  $\alpha^{-1}$  is also a square

This implies  $\beta$  is also a square

This is a contradiction to the assumption



c.

To prove: That in a finite field of even order, every element is a square

Consider;

$$\psi: K^* \rightarrow \{\pm 1\}$$

Such that;

$$\psi(x) = x^{(p^n-1)/2}$$

Since;

$$x^{p^n-1} = 1$$

As,  $K^*$  is a multiplicative group of order  $p^n - 1$

So, mapping to square roots of unity that is  $\{\pm 1\}$

Squares are certainly in the kernel of  $\psi$

On the other hand, since there are only  $(p^n - 1)/2$  solutions to

The squares are the only elements in the kernel.

Therefore,  $x$  is square if and only if  $\psi(x) = +1$

d.

To prove: that the irreducible polynomial for  $\gamma = \sqrt{2} + \sqrt{3}$  over  $\mathbb{Q}$  is reducible modulo  $p$  for every prime  $p$

First;

$$f = x^4 - 10x^2 + 1$$

This is the irreducible polynomial for  $\gamma$  since;

$$f(\gamma) = 0$$

$$\mathbb{Q}(\gamma) = \mathbb{Q}(\alpha, \beta)$$

And so;

$$[\mathbb{Q}(\gamma) : \mathbb{Q}] = 4$$

In this field following is the factorization;

$$f = (x - (\sqrt{2} + \sqrt{3}))(x - (\sqrt{2} - \sqrt{3}))(x - (-\sqrt{2} + \sqrt{3}))(x - (-\sqrt{2} - \sqrt{3}))$$

If  $f$  reduces over  $\mathbb{F}_p$ , then it must reduce to quadratic factors since;

$$\gamma \notin \mathbb{F}_p$$

Thus,  $f$  can factor in one of the three following ways by pairing up the factors above;

$$(x^2 - 1 - 2\sqrt{2})(x^2 - 1 + 2\sqrt{2}),$$

$$(x^2 + 1 - 2\sqrt{3})(x^2 + 1 + 2\sqrt{3}),$$

$$(x^2 - 5 - 2\sqrt{6})(x^2 - 5 + 2\sqrt{6})$$

The first factorization can occur if 2 is a square mod  $p$ , the second can occur if 3 is a square mod  $p$  and the last can occur if 6 is a square mod  $p$

So, at least one of 2, 3, 6 is a square in  $\mathbb{F}_p$ , that is at least one of these factorization is possible and  $f$  is reducible mod  $p$  for all  $p$

5. a

Order of a group is defined as the number of elements which is present in the given set or it can be defined as the cardinality of the set.

[Comment](#)

Step 2 of 4 ^

To prove: That any element of  $GL_2(\mathbb{Z})$  of finite order has order 1, 2, 3, 4, or 6

Let  $G$  be a finite subgroup of  $GL(2, \mathbb{Z})$

Then the mapping will be defined as;

$$G \cap SL(2, \mathbb{Z}) \rightarrow SL_2(\mathbb{F}_3)$$

This can be used to find all finite subgroups of  $GL(2, \mathbb{Z})$

The first step is to list all finite subgroups of  $SL(2, \mathbb{F}_3)$ , this has 24 elements.

Now, all the subgroups are having the orders 1, 2, 3, 4, 6

**a.**

By using the field theory;

Here, the order  $n$  of finite order elements in  $GL(2, \mathbb{Z})$  this has to satisfy;

$$\phi(n) \leq 2$$

Since, the minimal polynomial of an element of finite order must be divisible by the minimal polynomial over the rationals of a primitive  $n$ -th root of unity

Thus, the order will be of  $n = 1, 2, 3, 4, 6$

**b.**

By applying the Crystallographic Restriction;

Consider the matrix properties; the sum of the diagonal elements of a matrix is called the trace of the matrix. In two dimensional and three dimensional every rotation is planar rotation and the trace is a function of the angle alone.

The trace for two dimensional is  $2 \cos \theta$  and the trace for three dimensional is  $1 + 2 \cos \theta$

Now, consider the case of 6-fold, then the matrix will be;

$$\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

The trace is precisely 1, an integer

Thus further for different folds it can be said that the orders will be of  $n = 1, 2, 3, 4, 6$

6. a

A rational function is defined as an algebraic fraction such that both the numerator and the denominator are polynomials. The coefficients of the polynomials need not be rational and can be taken in any field such as  $K$

a.

To prove: that a rational function  $f(t)$  that generates the field  $\mathbb{C}(t)$  of all rational functions defines a bijective map  $T' \rightarrow T'$

Let

$$f(t) \in F \\ = \mathbb{C}(t)$$

And;

$$f = p/q$$

For some coprime  $p, q \in \mathbb{C}[t]$

Now,  $\mathbb{C}[t](f)$  is then isomorphic to  $F$  which is isomorphic to  $F[x]/(p-qx)$

Consider  $F[y]/(y)$  this is also isomorphic to  $F$

Now by using the proposition which states that;

Let  $f(t, x)$  and  $g(t, y)$  be irreducible polynomials in  $\mathbb{C}[t, x]$  and  $\mathbb{C}[t, y]$  respectively. Let

$K = F[x]/(f)$  and  $L = F[y]/(g)$  be the field extensions they define and let  $X$  and  $Y$  be the Riemann surfaces  $\{f=0\}$  and  $\{g=0\}$ . If  $K/F$  and  $L/F$  are isomorphic field extensions then  $X$  and  $Y$  are isomorphic branched coverings of  $T$

By using the above result the Riemann surfaces;

$$\{p-qx=0\}$$

And;

$$\{y=0\}$$

These are isomorphic branched coverings of  $T'$

Since,  $\{y=0\}$  is the complex  $t$ -plane  $T'$ , the isomorphism of coverings is an isomorphism  $T' \rightarrow T'$

b.

To prove: a rational function  $f(x)$  generates the field of rational functions  $\mathbb{C}(x)$  if and only if it is of the form  $(ax+b)/(cx+d)$  with  $ad-bc \neq 0$

Any function  $(ax+b)/(cx+d)$  with  $ad-bc \neq 0$  generates  $\mathbb{C}(x)$  by using the result which states that;

That every element of  $\mathbb{C}(x)$  can be written as a sum of a polynomial and a linear combination of functions of the form  $1/(x-a)^i$

By the above result any element in  $\mathbb{C}(x)$  can be expressed as a  $\mathbb{C}$ -linear combination of elements  $1/(cx+d)^i$  and then by using the Euclidean algorithm;

$$\frac{ax+b}{cx+d} = \frac{q(cx+d)+r}{cx+d} \\ = q + \frac{r}{cx+d}; q, r \in \mathbb{C}$$

This implies;  $\frac{1}{(cx+d)^i}$  for any  $i$  can be expressed as a  $\mathbb{C}$ -linear combination of  $\frac{(ax+b)^i}{(cx+d)^i}$

Conversely;

Suppose a rational function  $f(x)$  generates  $\mathbb{C}(x)$

But from the explanation of the above part it remains to show that an Automorphism of the Riemann surface  $T'$  are of the stated form.

Suppose;

$$f = \frac{p}{q}$$

This defines an Automorphism would not be bijective since then  $p$  has two zeroes that is, if  $q$  had degree larger than 1, considering  $\frac{1}{f}$  gives the same argument.

Further, one of  $p, q$  must be non-constant also to induce a bijection

Here,  $ad - bc \neq 0$  corresponds to having  $p, q$  coprime

c.

To identify: the group of automorphisms of  $\mathbb{C}(x)$  that are the identity on  $\mathbb{C}$

From the explanation of the above two parts  $\text{Aut}(\mathbb{C}(x))$  defined by;

$$x \rightarrow (ax+b)/(cx+d)$$

For  $ad - bc \neq 0$

This is the group under composition

Claim: There is a surjective homomorphism;

$$\phi: GL_2(\mathbb{C}) \rightarrow \text{Aut}(\mathbb{C}(x))$$

Defined by;

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow (ax+b)/(cx+d)$$

This map clearly maps;

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow x$$

This is also surjective and is a homomorphism since;

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} aa' + b'c & a'b + b'd \\ ac' + cd' & bc' + dd' \end{pmatrix}$$

Further both sides get mapped to;

$$\begin{aligned} \left( \frac{a'x+b'}{c'x+d'} \right) \circ \left( \frac{ax+b}{cx+d} \right) &= \frac{a'(ax+b) + b'(cx+d)}{c'(ax+b) + d'(cx+d)} \\ &= \frac{(aa' + b'c)x + (a'b + b'd)}{(ac' + cd')x + (bc' + dd')} \end{aligned}$$

It remains to find the kernel of  $\phi$

Also;

$$\begin{aligned} (ax+b)/(cx+d) &= x \\ &\in \mathbb{C}(x) \end{aligned}$$

If and only if;

$$ax+b = x(cx+d)$$

If and only if;

$$\begin{aligned} c &= b \\ &= 0 \end{aligned}$$

And;

$$a = d$$

Hence;

$\ker \varphi = \mathbb{C} \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  And so; by the first isomorphism theorem which states that;

If  $I \subset K$ , there is a unique homomorphism  $\bar{f}: \bar{R} \rightarrow R'$  such that  $\bar{f}\pi = f$

Then;

$$\begin{aligned} \text{Aut}(\mathbb{C}(x)) &\approx GL_2(\mathbb{C})/\ker \varphi \\ &= PGL_2(\mathbb{C}) \end{aligned}$$

The projective general linear group of order 2

## 7. a

Homomorphism is defined as the transformation of one set into another that preserves in the second set the relations between elements of the first.

To prove: that the homomorphism  $SL_2(\mathbb{Z}) \rightarrow SL_2(\mathbb{F}_p)$  obtained by reducing the matrix entries modulo  $p$  is surjective.

Let;

$$\pi: \mathbb{Z} \rightarrow \mathbb{F}_p$$

This be the quotient map reducing mod  $p$

First show that: if  $u, v \in \mathbb{F}_p$  are both zero, then there exist  $c, d \in \mathbb{Z}$  such that;

$$\pi(c) = u$$

$$\pi(d) = v$$

And;

$$\gcd(c, d) = 1$$

Now, suppose  $v \neq 0$  and let  $0 \leq \tilde{u}, \tilde{v} < p$  be the lifts of  $u, v$  in  $\mathbb{Z}$

And, since;

$$p \nmid \tilde{v}$$

Then there exists  $x, y \in \mathbb{Z}$  such that;

$$x\tilde{v} + yp = 1$$

Since,  $(p)$  is maximal in  $\mathbb{Z}$  and let;

$$c = x\tilde{u}\tilde{v} + yp$$

And;

$$d = \tilde{v}$$

Then;

$$\pi(c) = u$$

$$\pi(d) = v$$

And, since;

$$x\tilde{v} \equiv 1 \pmod{p}$$

Also;

$$\gcd(c, d) = 1$$

That is;

$$1 = cc + xd(1 - \tilde{u})$$

If  $v = 0$ , then necessarily  $u \neq 0$

And further reversing the role of  $u, v$  gives the required result.

Now, suppose that;

$$\begin{pmatrix} s & t \\ u & v \end{pmatrix} \in SL_2(\mathbb{F}_p)$$

Let  $0 \leq \tilde{s}, \tilde{t} < p$  be the lifts of  $s, t$  in  $\mathbb{Z}$

Then suppose  $c, d$  as constructed in the above part, than;

$$\tilde{s}d - \tilde{t}c = 1 + Np; N \in \mathbb{Z}$$

So, again suppose;

$$a = \tilde{s} + mp$$

$$b = \tilde{t} + np$$

Where  $m, n \in \mathbb{Z}$

Such that;

$$N = cn - dm$$

This is possible since;

$$\gcd(c, d) = 1$$

And;

$$\begin{aligned} ad - bc &= (\tilde{s} + mp)d - (\tilde{t} + np)c \\ &= \tilde{s}d - \tilde{t}c + (dm - cn)p \\ &= 1 + (N + dm - cn)p \\ &= 1 \end{aligned}$$

Hence;

$$\begin{pmatrix} s & t \\ u & v \end{pmatrix} \in SL_2(\mathbb{Z})$$

This map to;

$$\begin{pmatrix} s & t \\ u & v \end{pmatrix} \in SL_2(\mathbb{F}_p)$$

**Therefore, the homomorphism  $SL_2(\mathbb{Z}) \rightarrow SL_2(\mathbb{F}_p)$  is obtained by reducing the matrix entries modulo  $p$  is surjective.**



## Chapter 16

### Section 1

1. a

**Solution:** There are given some polynomials and we have to determine the orbit of the polynomials. Now if the polynomial is symmetric we will write it in terms of the elementary symmetric functions.

(a) The given polynomial is  $u_1^2 u_2 + u_2^2 u_3 + u_3^2 u_1$ .

The orbit consists of  $u_1^2 u_2 + u_2^2 u_3 + u_3^2 u_1$  and  $u_3^2 u_2 + u_2^2 u_1 + u_1^2 u_3$ , corresponding to odd and even permutations respectively in  $S_3$ .

(b) The given polynomial is  $(u_1 + u_2)(u_2 + u_3)(u_1 + u_3)$ .

Now notice that the given polynomial is clearly symmetric. Therefore the orbit consists of only the polynomial

$$(u_1 + u_2)(u_2 + u_3)(u_1 + u_3).$$

Now note that the polynomial is a homogeneous polynomial of degree 3.

Therefore we can write

$$(u_1 + u_2)(u_2 + u_3)(u_1 + u_3) = ax^3 + bxy + cz.$$

Let us now substitute the point  $(1, 0, 0)$ .

Then we have

$$2b = 2 \implies b = 1.$$

Let us now substitute the point  $(1, 1, 1)$ .

Then we have

$$c + 9 = 8 \implies c = -1.$$

Therefore the required symmetric function is

$$(u_1 + u_2)(u_2 + u_3)(u_1 + u_3) = xy - z.$$

(c) The given polynomial is  $(u_1 - u_2)(u_2 - u_3)(u_1 - u_3)$ .

The orbit consists of  $(u_1 - u_2)(u_2 - u_3)(u_1 - u_3)$  and  $(u_2 - u_1)(u_2 - u_3)(u_1 - u_3)$ , corresponding to odd and even permutations respectively in  $S_3$ .

(d) The given polynomial is  $u_1^3 u_2 + u_2^3 u_3 + u_3^3 u_1 - u_1 u_2^3 - u_2 u_3^3 - u_3 u_1^3$ .

The orbit consists of  $u_1^3 u_2 + u_2^3 u_3 + u_3^3 u_1 - u_1 u_2^3 - u_2 u_3^3 - u_3 u_1^3$  and

$u_3^3 u_2 + u_2^3 u_1 + u_1^3 u_3 - u_3 u_2^3 - u_2 u_1^3 - u_1 u_3^3$ , corresponding to odd and even permutations respectively in  $S_3$ .

(e) The given polynomial is  $u_1^3 + u_2^3 + \dots + u_n^3$ .

Now notice that the given polynomial is clearly symmetric. Therefore the orbit consists of only the polynomial  $u_1^3 + u_2^3 + \dots + u_n^3$ .

Now note that the polynomial is a homogeneous polynomial of degree 3.

Therefore we can write

$$u_1^3 + u_2^3 + \dots + u_n^3 = ax^3 + bxy + cz.$$

Let us now substitute the  $n$ -dimensional point  $(1, 0, 0, \dots, 0)$ .

Then we have  $a = 1$ .

Again substitute the  $n$ -dimensional point  $(1, 1, 0, \dots, 0)$  we have  $c = 3$ .

Therefore the required symmetric function is given by

$$u_1^3 + u_2^3 + \dots + u_n^3 = x^3 - 3xy + 3z.$$

This completes the solution.

## Result

3 of 3

Those are not symmetric just write down the orbit elements corresponding to odd and even permutations respectively in  $S_3$  and else write the symmetric function along with their orbits.

## 2. a

A set of symmetric polynomials  $\mathcal{S}$  is called as basis, if it satisfies the following conditions;

First is that any symmetric polynomial can be written as the sum of polynomials from  $\mathcal{S}$  with some coefficients

Second is that no polynomial from  $\mathcal{S}$  can be written as the sum of other polynomials from  $\mathcal{S}$

[Comment](#)

### Step 2 of 4 ^

To find: two basis for the ring of symmetric polynomial as a module over the ring  $R$

Let  $s_i = \sum x_1 \dots x_i$  be the fundamental symmetric polynomials

This symmetry can be written as;

$$\begin{aligned} k[s_1, \dots, s_n] &\subset k[s_1, \dots, s_n][x_1] \\ &\subset k[s_1, \dots, s_n][x_1] \\ &\vdots \\ &\subset k[s_1, \dots, s_n][x_1] \dots [x_n] \\ &= k[x_1, \dots, x_n] \end{aligned}$$

Consider the set of monomial polynomials as described below;

$$\{m_\lambda, \lambda = (\lambda_1 \geq \dots \geq \lambda_n \geq 0)\}$$

Further let  $a_\lambda = m_\lambda + \sum_{\mu < \lambda} a_{\lambda_\mu} m_\mu$

Where,

$$\{a_\lambda, \lambda = (n \geq \lambda_1 \geq \dots \geq \lambda_m \geq 0), m \in \mathbb{Z}_{>0}\}$$

Clearly  $a_\lambda$  satisfies both the condition for being a basis

Therefore, one of the basis of symmetric polynomials is  $m_\lambda$ , that is the set of monomial polynomials.

Next consider the power sum polynomials  $p_\lambda$ , which is defined as;

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \dots,$$

$$p_\lambda = m_\lambda, \text{ where } \lambda = (x, 0, \dots, 0)$$

Further expand the power sum polynomial as given below;

$$p_\lambda = a_\lambda m_\lambda + \sum_{\mu > \lambda} b_{\lambda\mu} m_\mu$$

Where,  $b_{\lambda\mu} \in \mathbb{Z}_{\geq 0}$

Here,  $a_\lambda$  is a natural number.

Thus, the given sum of power series of polynomial satisfies the conditions for being a basis.

Hence,  $p_\lambda$  is also a basis of symmetric polynomials

**Therefore, the two basis of symmetric polynomials is  $\boxed{m_\lambda \text{ and } p_\lambda}$ , that is the set of monomial polynomials.**

3. a

**Given:** We have  $w_k = u_1^k + u_2^k + \dots + u_n^k$ .

**To Prove:**

- (a) Newton's Identity:  $w_k - s_1 w_{k-1} + \dots \pm s_{k-1} w_1 \mp k s_k = 0$ .  
(b)  $w_1, \dots, w_n$  generate the ring of symmetric functions.

**Proof:**

- (a) Let us now consider the polynomial

$$f(x) = (x - u_1)(x - u_2) \dots (x - u_n).$$

Now notice that

$$f(x) = \sum_{i=0}^n (-1)^i s_i x^{n-i}.$$

Let us represent the derivative of  $f(x)$  as  $f'(x)$  and it yields that

$$\begin{aligned} f'(x) &= \sum_{i=1}^n \frac{f(x)}{x - u_i} \\ &= f(x) \cdot \sum_{i=1}^n \frac{1}{x - u_i} \\ &= f(x) \cdot \sum_{i=1}^n \left( \sum_{j=0}^{\infty} \frac{u_i^j}{x^{j+1}} \right) \\ &= f(x) \cdot \sum_{j=0}^{\infty} \frac{1}{x^{j+1}} \left( \sum_{i=1}^n u_i^j \right) \\ &= \left( \sum_{i=0}^n (-1)^i s_i x^{n-i} \right) \left( \sum_{j=0}^{\infty} \frac{w_j}{x^{j+1}} \right) \\ &= \left( \sum_{i=0}^{\infty} (-1)^i s_i x^{n-i} \right) \left( \sum_{j=0}^{\infty} \frac{w_j}{x^{j+1}} \right), \quad \text{since } s_i = 0 \quad \forall i > n \\ &= \sum_{i+j=l} (-1)^i s_i w_j x^{n-1-(i+j)} \\ &= \sum_{l=0}^{\infty} \left( \sum_{i=0}^{\infty} (-1)^i s_i w_{l-i} \right) x^{n-l-1}, \quad \text{since } w_{l-i} = 0 \quad \forall i > l \\ &= \sum_{l=0}^{\infty} (-1)^l s_l (n-l) x^{n-l-1}. \end{aligned}$$

So we have got the derivative  $f'(x)$  of the function  $f(x)$ .

Now notice that for each  $l \geq 0$ , we have

$$\sum_{i=0}^l (-1)^i s_i w_{l-i} = (-1)^l (n-l) s_l. \quad (1)$$

Now if we assume  $l \geq n$  then RHS of (1) will be zero and we have

$$\sum_{i=0}^l (-1)^i s_i w_{l-i} = 0.$$

Now if we assume  $l < n$ , then we have

$$\sum_{i=0}^l (-1)^i s_i w_{l-i} - (-1)^l (n-l) s_l = 0.$$

This completes the proof of Newton's Identity.

(b) Now look at the above Newton's Identity.

We conclude that the symmetric polynomials

$$s_n = \frac{1}{n} \left( \sum_{i=0}^{n-1} (-1)^{n+i} s_i w_{n-i} \right)$$

is defined by recursively via solely the given Newton sums  $w_k$ .

Note that the denominator in the recursion formula is an integer, therefore this definition is valid in the ring of symmetric functions.

This completes the solution.

## Result

3 of 3

First we prove the Newton's Identity by using derivative of  $f(x)$  and then using that we have solved that  $w_1, w_2, \dots, w_k$  generate the ring of symmetric functions.

## Section 2

1. a

**Solution:** We will propose to prove that the discriminant is a symmetric function.

Recall that the discriminant function is defined by

$$D(u) = (u_1 - u_2)^2(u_1 - u_3)^2 \dots (u_{n-1} - u_n)^2 = \prod_{i < j} (u_i - u_j)^2.$$

Let us consider that the variables  $u_i$  and  $u_j$  are swapped in the above equation  $D(u)$ .

Now notice that the factors of  $D(u)$  that does not contains  $u_i$  and  $u_j$  both are unaffected. And now the factors like  $(u_i - u_k)^2$  or  $(u_k - u_j)^2$ , where  $k \neq i, j$  transform into each other.

Now consider the factor  $(u_i - u_j)^2$ . After swapping it will become  $(u_j - u_i)^2$ , which is the same polynomial.

Therefore the discriminant remains unchanged under transpositions.

And notice that the transpositions generate all permutations.

This completes the proof.

## Result

2 of 2

Considering the discriminant function  $D(u)$  we have shown that swapping any two variable would not change the actual function.

## 2. a

**Solution:**

(a) We will propose to prove that the discriminant of a real cubic is non-negative if and only if the cubic has three real roots.

Notice that, since complex roots are always comes in pair, so a real cubic either has three real roots or one real root.

So there are two cases:

**Case-1:** It has three real roots.

In this case the discriminant will obviously be non-negative.

**Case-2:** It has one real root.

Let  $r$  be the real root and  $z, \bar{z}$  are the complex roots.

Note that the discriminant will be

$$D = (r - z)^2(r - \bar{z})^2(z - \bar{z})^2.$$

Now note that  $(z - \bar{z})$  is purely imaginary and non-zero, since  $z$  is not real.

Therefore  $(z - \bar{z})^2$  is a negative real number.

Now the product of the other two factors can be rewritten as

$$(r - z)^2(r - \bar{z})^2 = (r^2 - r(z + \bar{z}) + z\bar{z})^2.$$

Since  $z$  is non-real, notice that  $z + \bar{z}$  and  $z\bar{z}$  are both real.

Therefore  $(r - z)^2(r - \bar{z})^2$  is positive and since  $r \neq z$  and  $r \neq \bar{z}$ , so  $(r - z)^2(r - \bar{z})^2 \neq 0$ .

This proves that the discriminant is the product of a negative real and a positive real, and is therefore negative.



(b) Given that a real quartic polynomial has a positive discriminant.

Notice that, since complex roots are always comes in pair, a real quartic has zero, two, or four real roots.

By the hypothesis, since the polynomial has a positive discriminant, all four roots are distinct.

So there are two cases:

**Case-1:** It has four real roots.

In this case the discriminant will obviously be positive.

**Case-2:** It has exactly two real roots.

Let  $x, y$  are the real roots and  $z, \bar{z}$  are the complex roots.

Note that the discriminant will be

$$D = (x - z)^2(y - z)^2(x - \bar{z})^2(y - \bar{z})^2(z - \bar{z})^2(x - y)^2.$$

Now note that  $(z - \bar{z})$  is purely imaginary and non-zero, since  $z$  is not real.

Therefore  $(z - \bar{z})^2$  is a negative real number.

Now consider the product below as

$$(x - z)^2(x - \bar{z})^2 = (x^2 - x(z + \bar{z}) + z\bar{z})^2.$$

Since  $z$  is non-real, notice that  $z + \bar{z}$  and  $z\bar{z}$  are both real.

Therefore  $(x - z)^2(x - \bar{z})^2$  is positive and since  $x \neq z$  and  $x \neq \bar{z}$ , so  $(x - z)^2(x - \bar{z})^2 \neq 0$ .

Similarly notice that  $(y - z)^2(y - \bar{z})^2$  is positive and since  $y \neq z$  and  $y \neq \bar{z}$ , so  $(y - z)^2(y - \bar{z})^2 \neq 0$ .

Now the factor  $(x - y)^2$  is always positive.

So in this case the discriminant in this case is negative.

**Case-3:** It has no real roots.

Let  $z, \bar{z}$  and  $w, \bar{w}$  are all the complex roots.

Note that the discriminant will be

$$D = (z - \bar{z})^2(w - \bar{w})^2(z - w)^2(\bar{z} - \bar{w})^2(w - \bar{z})^2(z - \bar{w})^2.$$

Now notice that from Case-2, the factors  $(z - \bar{z})$  and  $(w - \bar{w})^2$  are each real and negative.

And the factors  $(z - w)^2, (\bar{z} - \bar{w})^2$  are complex conjugates of each other so their product is real and positive.

By the similar argument the product  $(w - \bar{z})^2(z - \bar{w})^2$  is real and positive. Thus, given the information that the discriminant is positive.

Therefore from above three cases it follows that the number of real roots of the quartic is either 0 or 4.

This completes the solution.

## Result

4 of 4

First we have shown that the discriminant of a real cubic is non-negative if and only if the cubic has three real roots and then we have proved that the number of real roots of a quartic is either 0 or 4, whose discriminant is positive.

### 3. a

(a)

Consider a cubic polynomial,

$$P(x) = x^3 - s_1x^2 + s_2x - s_3 \dots\dots (1)$$

With reference to the equation (16.2.5) it can say that the discriminant of (1) is given by

$$-4s_1^3s_3 + s_1^2s_2^2 + 18s_1s_2s_3 - 4s_2^3 - 27s_3^2 \dots\dots (2)$$

Consider a depressed cubic polynomial,

$$f(x) = x^3 + px + q \dots\dots (3)$$

With reference to the equation (16.2.8) it can say that the discriminant of the above equation is

$$-4p^3 - 27q^2 \dots\dots (4)$$



To transform the above cubic equation in depressed cubic equation  $f(x) = x^3 + px + q$

substitute  $x = y + \frac{s_1}{3}$  in (1),

$$P(x) = \left(y + \frac{s_1}{3}\right)^3 - s_1 \left(y + \frac{s_1}{3}\right)^2 + s_2 \left(y + \frac{s_1}{3}\right) - s_3$$

Simplify the above expression,

$$\begin{aligned} &= y^3 + \frac{s_1^3}{27} + 3y^2 \left(\frac{s_1}{3}\right) + 3y \left(\frac{s_1}{3}\right)^2 - s_1 \left(y^2 + \frac{s_1^2}{9} + \frac{2}{3}ys_1\right) + s_2y + \frac{s_2s_1}{3} - s_3 \\ &= y^3 + \frac{s_1^3}{27} + y^2s_1 + y\frac{s_1^2}{3} - s_1y^2 - \frac{s_1^3}{9} - \frac{2}{3}ys_1^2 + s_2y + \frac{s_2s_1}{3} - s_3 \\ &= y^3 + y \left(\frac{s_1^2}{3} + s_2 - \frac{2}{3}s_1^2\right) + \frac{s_1^3}{27} - \frac{s_1^3}{9} + \frac{s_2s_1}{3} - s_3 \quad \dots\dots (5) \\ &= y^3 + y \left(\frac{3s_2 - s_1^2}{3}\right) + \frac{-2s_1^3 + 9s_1s_2 - 27s_3}{27} \end{aligned}$$

Compare (3) and (5)

$$\begin{aligned} p &= \frac{3s_2 - s_1^2}{3} \\ q &= \frac{-2s_1^3 + 9s_1s_2 - 27s_3}{27} \quad \dots\dots (6) \end{aligned}$$

Substitute (6) in (4)

$$\begin{aligned} &-4 \left(\frac{3s_2 - s_1^2}{3}\right)^3 - 27 \left(\frac{-2s_1^3 + 9s_1s_2 - 27s_3}{27}\right)^2 \\ &= -\frac{4}{27} (3s_2 - s_1^2)^3 - \frac{1}{27} (-2s_1^3 + 9s_1s_2 - 27s_3)^2 \\ &= -\frac{4}{27} (-s_1^6 + 9s_2s_1^4 - 27s_2^2s_1^2 + 27s_2^3) \\ &= -\frac{1}{27} (4s_1^6 - 36s_2s_1^4 + 108s_3s_1^3 + 81s_2^2s_1^2 - 486s_2s_3s_1 + 729s_3^2) \\ &= -\frac{1}{27} \begin{pmatrix} -4s_1^6 + 36s_2s_1^4 - 108s_2^2s_1^2 + 108s_2^3 + 4s_1^6 \\ -36s_2s_1^4 + 108s_3s_1^3 + 81s_2^2s_1^2 - 486s_2s_3s_1 + 729s_3^2 \end{pmatrix} \\ &= -\frac{1}{27} (-27s_2^2s_1^2 + 108s_2^3 + 108s_3s_1^3 - 486s_2s_3s_1 + 729s_3^2) \\ &= -4s_1^3s_3 + s_1^2s_2^2 + 18s_1s_2s_3 - 4s_2^3 - 27s_3^2 \quad \dots\dots (7) \end{aligned}$$

The equation (7) is same as equation (2)

Thus it can say that the discriminant of cubic and depressed cubic equation are same that is Tschirnhausen substitution doesn't change the discriminant of the cubic polynomial.

**Hence Prove**

(b)

Consider a cubic polynomial,

$$P(x) = x^3 - s_1 x^2 + s_2 x - s_3$$

To transform the above cubic equation in depressed cubic equation

$$f(x) = x^3 + px + q \dots\dots (8)$$

Substitute  $x = y + \frac{s_1}{3}$  in the above equation,

$$P(x) = \left(y + \frac{s_1}{3}\right)^3 - s_1 \left(y + \frac{s_1}{3}\right)^2 + s_2 \left(y + \frac{s_1}{3}\right) - s_3$$

Simplify the above expression,

$$\begin{aligned} &= y^3 + \frac{s_1^3}{27} + 3y^2 \left(\frac{s_1}{3}\right) + 3y \left(\frac{s_1}{3}\right)^2 - s_1 \left(y^2 + \frac{s_1^2}{9} + \frac{2}{3} y s_1\right) + s_2 y + \frac{s_2 s_1}{3} - s_3 \\ &= y^3 + \frac{s_1^3}{27} + y^2 s_1 + y \frac{s_1^2}{3} - s_1 y^2 - \frac{s_1^3}{9} - \frac{2}{3} y s_1^2 + s_2 y + \frac{s_2 s_1}{3} - s_3 \\ &= y^3 + y \left(\frac{s_1^2}{3} + s_2 - \frac{2}{3} s_1^2\right) + \frac{s_1^3}{27} - \frac{s_1^3}{9} + \frac{s_2 s_1}{3} - s_3 \dots\dots (9) \\ &= y^3 + y \left(\frac{3s_2 - s_1^2}{3}\right) + \frac{-2s_1^3 + 9s_1 s_2 - 27s_3}{27} \end{aligned}$$

Comparing (8) and (9)

$$\begin{aligned} p &= \frac{3s_2 - s_1^2}{3} \\ q &= \frac{-2s_1^3 + 9s_1 s_2 - 27s_3}{27} \end{aligned}$$

Hence, the coefficients  $p$  and  $q$  that are obtain from the general cubic by the Tschirnhausen

substitution are  $\boxed{\begin{matrix} p = \frac{3s_2 - s_1^2}{3} \\ q = \frac{-2s_1^3 + 9s_1 s_2 - 27s_3}{27} \end{matrix}}.$

4. a

Consider the provided statement to determine the discriminant of the polynomial by using undetermined coefficients.

As it is known that from lemma 16.2.4, let  $D_n(p, q)$  be the determinant of  $x^n + px + q$  then the discriminant  $D_n$  of  $x^n + px + q, n \geq 2$  is given by,

$$D_n = (-1)^{\frac{(n-1)(n-2)}{2}} (n-1)^{n-1} p^n + (-1)^{\frac{n(n-1)}{2}} n^n q^{n-1}$$

[Comment](#)

Step 2 of 4 ^

(a)

Provided polynomial function is,

$$x^3 + px + q$$

As  $n > 2$  then the discriminant is provided as below;

$$\begin{aligned} D_3 &= (-1)^{\frac{(3-1)(3-2)}{2}} (3-1)^{3-1} p^3 + (-1)^{\frac{3(3-1)}{2}} 3^3 q^{3-1} \\ &= (-1)^{\frac{(2)(1)}{2}} (2)^2 p^3 + (-1)^{\frac{3(2)}{2}} (27) q^2 \\ &= (-1)(2)^2 p^3 + (-1)^1 (27) q^2 \\ &= -4p^3 - 27q^2 \end{aligned}$$

Hence, the discriminant of the polynomial is  $\boxed{-4p^3 - 27q^2}.$

(b)

Provided polynomial function is,

$$x^4 + px + q$$

As  $n > 2$  then the discriminant is provided as below;

$$\begin{aligned} D_4 &= (-1)^{\frac{(4-1)(4-2)}{2}} (4-1)^{4-1} p^4 + (-1)^{\frac{4(4-1)}{2}} 4^4 q^{4-1} \\ &= (-1)^{\frac{(3)(2)}{2}} (3)^3 p^4 + (-1)^{\frac{4(3)}{2}} (256) q^3 \\ &= (-1)^{(3)} (3)^3 p^4 + (-1)^6 (256) q^3 \\ &= -27 p^4 + 256 q^3 \end{aligned}$$

Hence, the discriminant of the polynomial is  $D_4 = \boxed{-27 p^4 + 256 q^3}$ .

(c)

Provided polynomial function is,

$$x^5 + px + q$$

As  $n > 2$  then the discriminant is provided as below;

$$\begin{aligned} D_5 &= (-1)^{\frac{(5-1)(5-2)}{2}} (5-1)^{5-1} p^5 + (-1)^{\frac{5(5-1)}{2}} 5^5 q^{5-1} \\ &= (-1)^{\frac{(4)(3)}{2}} (4)^4 p^5 + (-1)^{\frac{5(4)}{2}} (3125) q^4 \\ &= (-1)^6 (4)^4 p^5 + (-1)^{10} (3125) q^4 \\ &= 256 p^5 + 3125 q^4 \end{aligned}$$

Hence, the discriminant of the polynomial is  $D_5 = \boxed{256 p^5 + 3125 q^4}$ .

5. a

For a polynomial  $P(x)$  of degree  $n$  with roots  $u_1, u_2, \dots, u_n$  which is given as follows

$$P(x) = x^n - s_1 x^{n-1} + s_2 x^{n-2} - \dots \pm s_n,$$

Where  $s_i$  denote the elementary symmetric function.

For such polynomial function the discriminant is denoted by  $D(u)$  and is defined as

$$D(u) = \prod_{i < j} (u_i - u_j)^2$$

Consider the polynomial  $P(x)$  of degree 4 given by

$$P(x) = x^4 - s_1x^3 + s_2x^2 - s_3x + s_4$$

Let  $u_1, u_2, u_3, u_4$  be the roots of the above defined polynomial

Then discriminant of  $P(x)$  is given by

$$D(u) = \prod_{i < j} (u_i - u_j)^2 \\ = (u_1 - u_2)^2 (u_1 - u_3)^2 (u_1 - u_4)^2 (u_2 - u_3)^2 (u_2 - u_4)^2 (u_3 - u_4)^2$$

Use the following transformation in order to evaluate the discriminant in the terms of  $s_1, s_2, s_3, s_4$

$$s_1 = u_1 + u_2 + u_3 + u_4 \\ s_2 = u_1u_2 + u_1u_3 + u_1u_4 + u_2u_3 + u_2u_4 + u_3u_4 \\ s_3 = u_1u_2u_3 + u_1u_2u_4 + u_1u_3u_4 + u_2u_3u_4 \\ s_4 = u_1u_2u_3u_4$$

Then the discriminant  $\Delta(s_1, s_2, s_3, s_4)$  is given by

$$\Delta(s_1, s_2, s_3, s_4) = (s_1^2s_2^2s_1^2 - 4s_3^3s_1^3 - 4s_2^2s_2^3 + 18s_3^3s_2s_1 - 27s_3^4 + 256s_4^3) + s_4(q) + s_4^2(p) \\ = r + s_4(q) + s_4^2(p)$$

are expressions obtained from  $s_1, s_2, s_3, s_4$

Clearly the term  $s_4(q)$  and  $s_4^2(p)$  are divisible by  $s_4$

So, the only terms in the discriminant which are not divisible by  $s_4$  are coming from the expansion of  $r$  only.

Since  $r = s_3^2s_2^2s_1^2 - 4s_3^3s_1^3 - 4s_2^2s_2^3 + 18s_3^3s_2s_1 - 27s_3^4 + 256s_4^3$ , then the last term is divisible by  $s_4$ .

So the only terms which are not divisible by  $s_4$  are as follows

$$s_3^2s_2^2s_1^2, 4s_3^3s_1^3, 4s_2^2s_2^3, 18s_3^3s_2s_1 \text{ And } 27s_3^4$$

So the coefficients of the terms which are not divisible by  $s_4$  are given as follows

$$s_3^2s_2^2s_1^2 \rightarrow 1 \\ s_3^3s_1^3 \rightarrow 4 \\ s_2^2s_2^3 \rightarrow 4 \\ s_3^3s_2s_1 \rightarrow 18$$

**Therefore the coefficients in  $\Delta(s_1, s_2, s_3, s_4)$  of all monomials not divisible by  $s_4$  are 1, 4, 4 and 18 .**

6. a

For a polynomial  $P(x)$  of degree  $n$  with roots  $u_1, u_2, \dots, u_n$  which is given as follows

$$P(x) = x^n - s_1x^{n-1} + s_2x^{n-2} - \dots \pm s_n$$

Where  $s_i$  denote the elementary symmetric function.

For such polynomial function the discriminant is denoted by  $D(u)$  and is defined as

$$D(u) = \prod_{i < j} (u_i - u_j)^2$$

Let  $u'_i = u_i + t$  for  $i = 1, 2, 3$

Then the elementary symmetric functions are denoted by  $s'_1, s'_2, s'_3$  and defined by

$$s'_1 = \sum_{i=1}^3 u'_i(t)$$

$$s'_2 = u'_1 u'_2 + u'_2 u'_3 + u'_1 u'_3$$

$$s'_3 = u'_1 u'_2 u'_3$$

Now differentiate each  $s'_i$  with respect to  $t$

For  $i = 1$ ,

$$\begin{aligned} \frac{d(s'_1)}{dt} &= \frac{d\left(\sum_{i=1}^3 u'_i(t)\right)}{dt} \\ &= \frac{du'_1}{dt} + \frac{du'_2}{dt} + \frac{du'_3}{dt} \\ &= 1 + 1 + 1 \\ &= 3 \end{aligned}$$

For  $i = 2$ ,

$$\begin{aligned} \frac{d(s'_2)}{dt} &= \frac{d(u'_1 u'_2 + u'_2 u'_3 + u'_1 u'_3)}{dt} \\ &= \frac{d(u'_1 u'_2)}{dt} + \frac{d(u'_2 u'_3)}{dt} + \frac{d(u'_1 u'_3)}{dt} \\ &= 2(u'_1 + u'_2 + u'_3) \\ &= 2s'_1 \end{aligned}$$

For  $i = 3$ ,

$$\begin{aligned} \frac{d(s'_3)}{dt} &= \frac{d(u'_1 u'_2 u'_3)}{dt} \\ &= \frac{d(u'_1)}{dt} u'_2 u'_3 + u'_1 \frac{d(u'_2 u'_3)}{dt} \\ &= u'_1 u'_2 + u'_2 u'_3 + u'_1 u'_3 \\ &= s'_2 \end{aligned}$$

Now evaluate  $\Delta(u')$

$$\begin{aligned} \Delta(u') &= (u'_1 - u'_2)(u'_1 - u'_3)(u'_2 - u'_3) \\ &= (u_1 + t - u_2 - t)(u_1 + t - u_3 - t)(u_2 + t - u_3 - t) \\ &= (u_1 - u_2)(u_1 - u_3)(u_2 - u_3) \end{aligned}$$

Now differentiate both sides with respect to  $t$

$$\begin{aligned} \frac{d(\Delta(u'))}{dt} &= \frac{d((u_1 - u_2)(u_1 - u_3)(u_2 - u_3))}{dt} \\ &= 0 \end{aligned}$$

Now since  $\Delta(u') = \Delta(u)$ ,

Using above calculated values and put them in the discriminant of a cubic equation the result holds good.

That is, the discriminant of cubic equation comes out to be as follows

$$\Delta(u) = -4u_1^3 u_3 + u_1^2 u_2^2 + 18u_1 u_2 u_3 - 4u_2^3 - 27u_3^2$$

**Therefore, the result stated in the question has been proved.**

7. a

Consider the provided statement to prove that  $(p-q)^2 = D(u)$ . As it is provided that, there are  $n$  variables such that  $m = u_1 u_2^2 u_3^3 \dots u_{n-1}^{n-1}$  and  $p(u) = \sum_{\sigma \in A_n} \sigma(m)$ .

Comment

Step 2 of 3 ^

Let  $R[u_1, \dots, u_n]$  be the polynomial ring and let  $\delta(u) = \prod_{i < j} (u_i - u_j)$ . It is claim that  $u_i - u_j \mid p - q$  for all  $i < j$ . This is equivalent to showing that  $p - q = 0$  if  $u_i - u_j = 0$ . Therefore it is assume that  $u_i = u_j$  for  $i < j$  and let  $\tau = (ij) \in S_n$ . Then,  $p(u) = p(\tau(u))$  and  $q(u) = q(\tau(u))$  then,

$$\begin{aligned} p(\tau(u)) &= \sum_{\sigma \in A_n} \sigma(\tau(m)) \\ &= \sum_{\sigma \in A_n, \tau} \sigma'(m) \\ &= q(u) \end{aligned}$$

Similarly,  $p(\tau(u)) = p(u)$ , since  $A_n$  has exactly two cosets in  $S_n$ .

Therefore,  $p(u) - q(u) = q(u) - p(u)$  and  $p(u) - q(u) = 0$ .

Since,  $\delta(u)$  is homogeneous of degree  $\frac{n(n-1)}{2}$  as are  $p, q, u_i - u_j \mid p - q, \forall i < j$ . It implies that,

$$p - q = a \prod_{i < j} (u_i - u_j), a \in R$$

As  $a = \pm 1$  because in the equation  $p - q = a \prod_{\sigma \in S_n} \text{sgn}(\sigma) \sigma(m)$

The coefficient on  $m$  is 1, while in the expression  $\delta(u) = \prod_{i < j} (u_i - u_j)$  the coefficient on  $m$  is  $\pm 1$ . Therefore,

$$\begin{aligned} (p - q)^2 &= \delta(u)^2 \\ &= D(u) \end{aligned}$$

Hence, provided statement is **proved**.

## Section 3

1. a

**Given:**  $f$  is a polynomial of degree  $n$  with coefficients in  $F$  and  $K$  is the splitting field for  $f$  over  $F$ .

**To Prove:**  $[K : F]$  is a divisor of  $n!$ .

**Proof:** We will propose to prove that  $[K : F]$  divides  $n!$  by using induction on the degree of  $f$ .

Now notice that if the degree is 1 then the proposal is trivial.

Now assume the proposal is true for all degrees smaller than  $n$  and consider that  $f$  has degree  $n$ . We will solve it in two steps.



**Step-1:** Let us assume that  $f$  is irreducible.

Let  $a \in K$  is a root of  $f$ . Then we have

$$f(x) = (x - a)g(x).$$

Clearly notice that

$$[F(a) : F] = n \quad \text{and} \quad g(x) \in F(a)[x].$$

Now notice that there exists a field  $L$  satisfying

$$F(a) \subseteq L \subseteq K$$

which is a splitting field of  $g$ , since  $g$  splits completely in  $K$  and by the inductive hypothesis, we have

$$[L : F(a)] \text{ divides } (n - 1)!.$$

Now note that since  $L$  contains all the roots of  $g$  along with  $a$ , it contains all the roots of  $f$ .

Therefore we have

$$K \subseteq L.$$

This follows that

$$K = L.$$

Now we have

$$\begin{aligned} [K : F] &= [K : F(a)][F(a) : F] \\ &= n \times [K : F(a)] \\ &= n \times [L : F(a)]. \end{aligned}$$

Since  $[L : F(a)]$  divides  $(n - 1)!$  we have  $[K : F]$  divides  $n(n - 1)!$ , that is  $[K : F]$  divides  $n!$ .

Step-1 is completed.

**Step-2:** Let us assume the  $f$  is not irreducible.

Then  $f$  can be written as

$$f = gh, \quad \text{where } g \text{ and } h \text{ both have positive degree.}$$

Let us consider  $H$  be a subfield of  $K$  obtained by adjoining the roots of  $g$  to  $F$ . Then notice that  $H$  is a splitting field for  $g$ .

Now by induction hypothesis we have

$$[H : F] \text{ divides } m!, \quad \text{where } m = \deg(g).$$

Now adjoining the roots of  $h$  to  $H$  must yield  $K$ .

Therefore so  $K$  is a splitting field for  $h$  when considered as a polynomial over  $H$ .

Now the induction hypothesis implies

$$[K : H] \text{ divides } t!, \quad \text{where } t = \deg(h).$$

Therefore we have

$$\begin{aligned} [K : F] &= [K : H][H : F] \\ \implies [K : F] &\text{ divides } (m!) \times (t!) \\ \implies [K : F] &\text{ divides } (m + t)! = n!. \end{aligned}$$

This follows that  $[K : F]$  divides  $(n!)$ .

This completes the proof.

First we consider  $f$  is irreducible and proved the proposal and then proved without assuming that  $f$  is irreducible.

2. a

(a)

Consider the following polynomial:

$$x^3 - 2$$

To find the root from the above polynomial, then

$$\alpha^3 - 2 = 0$$

$$\alpha^3 = 2$$

$$\alpha = 2^{1/3}$$

$$\alpha = \sqrt[3]{2}$$

Thus, a root of  $x^3 - 2$  is  $\alpha = \sqrt[3]{2}$ .

Then after factoring and applying quadratic formula,

$$x^3 - 2 = (x - \alpha)(x - \alpha\omega)(x - \alpha\omega^2)$$

Where,  $\omega$  is a complex cube root of unity.

$$\omega^2 + \omega + 1 = 0$$

And  $\omega \notin \mathbb{R}$  hence  $\omega \notin \mathbb{Q}(\alpha)$

So, the splitting field of  $x^3 - 2$  has degree  $3 \times 2 = 6$  in fact the splitting field is  $\mathbb{Q}(\alpha, \omega)$ .

(b)

Consider the following polynomial:

$$x^4 - 1$$

To find the root from the above polynomial, then

$$x^4 - 1 = 0$$

$$x^4 = 1$$

The above function  $x^4 = 1$  means that the fourth root of unity is the roots of the polynomials.

Then after factoring and applying quadratic formula,

$$\begin{aligned} x^4 - 1 &= (x^2 + 1)(x + 1)(x - 1) \\ &= (x + i)(x - i)(x + 1)(x - 1) \end{aligned}$$

So the splitting field is  $\mathbb{Q}(i)$  which has degree 2 over  $\mathbb{Q}$  since  $i$  satisfies the irreducible polynomial  $x^2 + 1$ .

Therefore, the function  $x^4 - 1$  has the degree of the splitting fields is 2.

(c)

Consider the following polynomial:

$$x^4 + 1$$

To find the root from the above polynomial, then

$$x^4 + 1 = 0$$

$$x^4 = -1$$

$$x^2 = \pm i$$

$$x = \pm \sqrt{\pm i}$$

The above function  $x^4 = -1$  means that the fourth root of unity is the roots of the polynomials.

The splitting field will contain both  $i$  and  $\sqrt{2}$  so, claim that the degree of the splitting field is 4.

Therefore, the function  $x^4 + 1$  has the degree of the splitting fields is 4.

### 3. a

Let  $F = \mathbb{F}_2(u)$  be the field of rational functions over the prime field  $\mathbb{F}_2$ .

To prove the polynomial  $x^2 - u$  is irreducible over  $F$ , and that it has a double root in a splitting field.

[Comment](#)

#### Step 2 of 4 ^

First show that the polynomial  $x^2 - u$  has a double root in a splitting field.

Let  $\alpha$  be a root of the polynomial in a splitting field, that is

$$x^2 - u = 0$$

$$\alpha^2 - u = 0$$

$$\alpha^2 = u$$

And note that in characteristic 2,

Then,

$$\begin{aligned} (x - \alpha)^2 &= x^2 - \alpha^2 \\ &= x^2 - u \end{aligned}$$

Therefore,  $\alpha$  is a double root in a splitting field.

Show that the polynomial  $x^2 - u$  is irreducible over  $F$ .

Suppose  $x^2 - u$  is not irreducible, then it is reducible and it would have a root - some rational function  $f(u)$  as the form of  $\frac{p(u)}{q(u)}$ .

However, the degree of  $f(u)$  computed as the difference of degrees of the numerator and denominator, then have to  $\frac{1}{2}$  which is not possible because degree of polynomial is always whole numbers.

So, assumption was wrong.

Therefore,  $x^2 - u$  is irreducible.

**Hence proved**

3.2 Determine the degrees of the splitting fields of the following polynomials over  $\mathbb{Q}$

(a)  $x^3 - 2$

Solution;

The three roots of  $x^3 - 2$  are  $x^3 - 2 = 0$

Solving we get three roots of  $x^3 - 2$  are  $2^{1/3}, 2^{1/3}(\frac{-1 + \sqrt{3}}{2}), 2^{1/3}(\frac{-1 - \sqrt{3}}{2})$ ,

The smallest field containing  $\mathbb{Q}$  and the above roots is same as the splitting field of  $x^3 - 2$  over  $\mathbb{Q}$  is  $\mathbb{Q}(2^{1/3}, \sqrt{3}i)$  so  $\deg[\mathbb{Q}(2^{1/3}) : \mathbb{Q}] = \deg(x^3 - 2) = 3$

Again  $x^2 + 3 \in \mathbb{Q}[x]$  is the irreducible polynomial over  $\mathbb{Q} \subset \mathbb{Q}(2^{1/3})$  is of lowest degree which is satisfied by  $\sqrt{3}i$

Hence  $[\mathbb{Q}(2^{1/3}, \sqrt{3}i) : \mathbb{Q}] = 3 \times 2 = 6$

(b)  $x^4 - 1$

Solution:  $x^4 - 1 = (x^2 - 1)(x^2 + 1) = 9x - 1)(x + 1)(x - i)(x + i)$

The roots are  $\pm 1, \pm i$  where  $i = \sqrt{-1}$

Hence the splitting field of  $x^4 - 1$  over  $A$  are  $\mathbb{Q}(i)$  and  $[\mathbb{Q}(i) : \mathbb{Q}] = 2$  since  $x^2 + 1$  is the irreducible polynomial of degree 2 over  $\mathbb{Q}$  which is satisfied by  $i$ .

(c) the roots of  $x^4 + 1$  are given by

$$\frac{1}{\sqrt{2}}(1+i), -\frac{1}{\sqrt{2}}(1-i), -\frac{1}{\sqrt{2}}(1+i), \frac{1}{\sqrt{2}}(1-i)$$

Splitting field  $K$  of  $x^4 + 1$  over  $\mathbb{Q}$  is  $K = \mathbb{Q}(\pm \frac{1}{\sqrt{2}}(1+i), \pm \frac{1}{\sqrt{2}}(1-i))$

We show  $K = \mathbb{Q}(\sqrt{2}, i)$

$$\Rightarrow \frac{1 \pm i}{\sqrt{2}} \in K$$

$$\Rightarrow \sqrt{2}, i \in K$$

Also  $\mathbb{Q}$

$$\subset K \text{ and thus } \mathbb{Q}(\sqrt{2}, i) \subseteq K$$

$$\text{Again } \sqrt{2}, i \in \mathbb{Q}(\sqrt{2}, i)$$

$$\Rightarrow \pm \frac{1}{\sqrt{2}}(1 \pm i) \in \mathbb{Q}(\sqrt{2}, i)$$

$$\Rightarrow K \subseteq \mathbb{Q}(\sqrt{2}, i) \text{ as } \mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}, i)$$

$$K = \mathbb{Q}(\sqrt{2}, i)$$

now

Now  $x^2 + 1 \in \mathbb{Q}(\sqrt{2})[x]$  is irreducible over  $\mathbb{Q}(\sqrt{2})$

$$\Rightarrow [\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}(\sqrt{2})] = \deg \text{lrr}(\mathbb{Q}(\sqrt{2}, i)) = 2$$

As  $i$  satisfies  $x^2 + 1$

$$[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = \deg \text{lrr}(\mathbb{Q}, \sqrt{2}) = \deg(x^2 - 2) = 2$$

$$[K : \mathbb{Q}] = [K : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \times 2 = 4$$

!

Let  $F = F_2(u)$  be the rational function field over the field of two elements. Prove that the

polynomial  $x^2 - u$  is irreducible in  $F[x]$  and that it has two equal roots in a splitting field. Solution:

First, let  $\alpha$  be a root of the polynomial in a splitting field, and note that in characteristic 2,  $(x - \alpha)^2 = x^2 - \alpha^2 = x^2 - u$ . Therefore,  $\alpha$  is a double root. Now, suppose  $x^2 - u$  weren't irreducible, then it would have a root - some rational function  $f(u)$ . However, the degree of  $f(u)$  (computed as the difference of degrees of the numerator and denominator) would then have to be  $1/2$ , which doesn't happen.

## Section 4

1. a

(a)

Consider the provided statement to determine all the automorphism of the given field.

As provided field is  $\mathbb{Q}(\sqrt[3]{2})$  and  $\mathbb{Q}(\sqrt[3]{2}, \omega)$ ,  $\omega = e^{\frac{2\pi i}{3}}$

It is known that any automorphism  $\varphi$  of field must leave  $\mathbb{Q}$  fixed. As from the previous exercise, it is obtained that  $\varphi(1) = 1$ . As multiplicative inverses are preserved through  $\varphi$  then any rational which is in the form of  $\frac{p}{q}$  must be fixed.

It is assumed that  $\varphi \in \text{Aut}\mathbb{Q}(\sqrt[3]{2})$ . Then  $\varphi$  is determined by  $\varphi(\sqrt[3]{2})$  and it is also necessary that  $\varphi(\sqrt[3]{2})^3 = 2$ . Therefore,

$$\varphi(\sqrt[3]{2}) = \omega^j \sqrt[3]{2}, j \in \{0, 1, 2\}$$

Then,  $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$  and  $\omega \in \frac{\mathbb{C}}{\mathbb{R}}$  hence  $j = 0$  and  $\varphi = id$ .

Now it is assumed that  $\varphi \in \text{Aut}\mathbb{Q}(\sqrt[3]{2}, \omega)$  then  $\varphi$  is determined by  $\varphi(\sqrt[3]{2})$  and  $\varphi(\omega)$ .

It is necessary that  $\varphi(\sqrt[3]{2})^3 = 2$  and  $\varphi(\omega)^3 = 1$ .

Therefore,

$$\varphi(\sqrt[3]{2}) = \omega^j \sqrt[3]{2}$$

$$\varphi(\omega) = \omega^k$$

Where the value of  $j, k \in \{0, 1, 2\}, k \neq 0$ , otherwise  $\varphi$  is not surjective.

Hence,  $\mathbb{Q}(\sqrt[3]{2}, \omega)$  has  $\boxed{6}$  automorphisms.

(b)

To find all automorphisms of  $K$  for the provided polynomial function,

$$\begin{aligned} f(x) &= (x^2 - 2x - 1)(x^2 - 2x - 7) \\ &= x^4 - 2x^3 - 7x^2 - 2x^3 + 4x^2 + 14x - x^2 + 2x + 7 \\ &= x^4 - 6x^3 - 4x^2 + 16x + 7 \end{aligned}$$

The roots of the equation  $(x^2 - 2x - 7)$  is calculated as below,

$$\begin{aligned} x &= \frac{-(-2) \pm \sqrt{(-2)^2 - 4 \cdot 1 \cdot (-7)}}{2 \cdot 1} \\ &= \frac{2 \pm \sqrt{4 + 28}}{2} \\ &= \frac{2 \pm 4\sqrt{2}}{2} \\ &= 1 \pm 2\sqrt{2} \end{aligned}$$

The roots of the equation  $(x^2 - 2x - 1)$  is calculated as below,

$$\begin{aligned} x &= \frac{-(-2) \pm \sqrt{(-2)^2 - 4 \cdot 1 \cdot (-1)}}{2 \cdot 1} \\ &= \frac{2 \pm \sqrt{4 + 4}}{2} \\ &= \frac{2 \pm 2\sqrt{2}}{2} \\ &= 1 \pm \sqrt{2} \end{aligned}$$

Therefore, the roots of the equation are  $1 \pm \sqrt{2}, 1 \pm 2\sqrt{2}$ . Hence,  $K = \mathbb{Q}(\sqrt{2})$  have only the trivial automorphism and the automorphism sending from  $\sqrt{2}$  to  $-\sqrt{2}$ .

Hence, the number of automorphism is  $\boxed{2}$ .



## Section 5

1. a

**Solution:** We will determine the group of automorphisms that following sets of automorphisms of the field of rational functions  $\mathbb{C}(t)$  and the fixed field.

$$(a) \sigma(t) = t^{-1}$$

Note that The automorphism  $\sigma$  here is an involution.

Therefore the group it generates consists of only itself and the identity element.

Let us now consider  $K = \mathbb{C}(t + t^{-1})$ .

Notice that  $K$  is a sub-field of the fixed field. Now observe that  $t$  is a root of the polynomial  $f(x)$ , which has coefficients in  $K$ , where

$$\begin{aligned} f(x) &= (x - t)(x - t^{-1}) \\ &= x^2 - (t + t^{-1})x + 1. \end{aligned}$$

Therefore the degree of  $[\mathbb{C}(t) : K]$  is at most 2.

Now by **Fixed Field Theorem** it follows that  $K$  itself is the fixed field, since  $K$  is a sub-field of the fixed field and the fixed field has degree 2.

$$(b) \sigma(t) = it$$

Notice that the orbit of  $t$  under  $\sigma$  is  $t, it, -t, -it, t, \dots$

Therefore,  $\sigma$  generates a cyclic group of order 4.

Let us consider now  $K = \mathbb{C}(t^4)$ .

Now note that

$$[\mathbb{C}(t) : K] \leq 4.$$

Therefore the degree  $[\mathbb{C}(t) : K]$  is at most 4 and  $K$  is a subfield of the fixed field, over which  $\mathbb{C}(t)$  has degree 4.

This follows that  $K$  is itself the fixed field.

$$(c) \sigma(t) = -t \text{ and } \tau(t) = t^{-1}$$

First notice that both  $\sigma$  and  $\tau$  are involutions. Therefore the automorphism group is the Klein four-group, 4 order abelian group with all non-zero elements has order 2.

We will propose to prove that the fixed field is  $K = \mathbb{C}(t^2 + t^{-2})$ .

This follows from the fact that  $K$  is fixed by both  $\sigma$  and  $\tau$  and note that  $t$  is a root of the polynomial  $f(x)$  where

$$f(x) = x^4 - (t^2 + t^{-2})x^2 + 1.$$

Now note that

$$[\mathbb{C}(t) : K] \leq 4.$$

Therefore the degree  $[\mathbb{C}(t) : K]$  is at most 4.

$$(d) \sigma(t) = wt \text{ and } \tau(t) = t^{-1}, \text{ where } w = e^{2\pi i/3}$$

Now notice that  $\sigma$  has order 3 and  $\tau$  has order 2.

Now we have

$$\begin{aligned} \sigma(\tau(t)) &= w^{-1}t^{-1} \\ &= w^2t^{-1}, \text{ since } w^3 = 1 \\ &= \tau(\sigma^2(t)). \end{aligned}$$

It follows that the automorphism group is the symmetric group  $S_3$  and note that the degree of  $\mathbb{C}(t)$  over the fixed field is 6.



We will propose to prove that the fixed field is  $K = \mathbb{C}(t^3 + t^{-3})$ .

This follows from the fact that  $K$  is fixed by both  $\sigma$  and  $\tau$  and note that  $t$  is a root of the polynomial  $f(x)$  where

$$f(x) = x^6 - (t^3 + t^{-3})x^3 + 1.$$

Now note that

$$[\mathbb{C}(t) : K] \leq 6.$$

Therefore the degree  $[\mathbb{C}(t) : K]$  is at most 6.

This completes the solution.

## Result

3 of 3

We have used Fixed Field Theorem and the given automorphisms to determine the group of automorphisms that following sets of automorphisms of the field of rational functions  $\mathbb{C}(t)$  and the fixed field.

## 2. a

Field is defined as the nonzero commutative division ring whose nonzero elements form an abelian group under multiplication

To show: that the automorphisms  $\sigma(t) = \frac{t+i}{t-i}$  and  $\tau(t) = \frac{it-i}{t+1}$  of  $\mathbb{C}(t)$  generate a group isomorphic to the alternating group  $A_4$ , and determining the fixed field of this group.

First calculate the products of  $\sigma, \tau$  by taking  $id = t$ :

$$\begin{aligned}\sigma\tau &= -t \\ \sigma\tau^2\sigma &= \frac{1}{t} \\ \tau\sigma &= -\frac{1}{t} \\ \sigma &= \frac{t+i}{t-i}\end{aligned}$$

Further;

$$\begin{aligned}\sigma^2\tau &= \frac{t-i}{t+i} \\ \tau\sigma^2 &= \frac{it+1}{-it+1} \\ \tau^2 &= \frac{-it+1}{it+1} \\ \sigma^2 &= \frac{t+1}{-it+i}\end{aligned}$$

And;

$$\begin{aligned}\tau &= \frac{it-i}{t+1} \\ \tau^2\sigma &= \frac{t+1}{it-i} \\ \sigma\tau^2 &= \frac{-it+1}{t+1}\end{aligned}$$

Here, use the relation;

$$\begin{aligned}\sigma^3 &= \tau^3 \\ &= \sigma\tau\sigma\tau \\ &= id\end{aligned}$$

Thus, there are 12 elements in this group  $G$ .

Consider the four pairs of 3 points defining circles in  $\mathbb{CP}^1$ :

$$\begin{aligned}&\{\{0, -i, 1\}, \{-1, i, \infty\}\} \\ &\{\{0, -1, i\}, \{-i, 1, \infty\}\} \\ &\{\{0, i, 1\}, \{-1, -i, \infty\}\} \\ &\{\{0, -1, i\}, \{1, i, \infty\}\}\end{aligned}$$

Now, label these four pairs as 1, 2, 3, 4 respectively

And,  $\sigma$  permutes these pairs as (123) and  $\tau$  as (234)

Hence,  $G$  is isomorphic to a subgroup of  $S_4$

This means it has an order of 12

Further by using the result which states that;

The only subgroup of order 12 of the symmetric group  $S_4$  is alternating group  $A_4$

Thus,  $G \cong A_4$

To find the fixed field, put the subgroup of  $A_4$  isomorphic to the Klein 4-group  $C_2 \times C_2$

Further by using the result of the theorem which states that;

Let  $H$  be a finite group of automorphisms of a field  $K$  and let  $F$  denote the fixed field  $K^H$ . Let  $\beta_1$  be an element of  $K$ , and let  $\{\beta_1, \dots, \beta_r\}$  be the  $H$ -orbit of  $\beta_1$ , then the irreducible polynomial for  $\beta_1$  over  $F$  is;

$$\begin{aligned}g(x) &= (x - \beta_1) \dots (x - \beta_r) \\ &= x^r - b_1 x^{r-1} + \dots \pm b_r\end{aligned}$$

By using the above result the irreducible polynomial for  $t$  over the fixed field is the polynomial whose roots form its orbit;

$$\begin{aligned}(x-t)(x+t)(x-1/t)(x+1/t) &= (x^2 - t^2)(x^2 - 1/t^2) \\ &= x^4 - (t^2 + t^{-2})x^2 + 1\end{aligned}$$

Now, letting;

$$u = t^2 + t^{-2}$$

Then;

$$[\mathbb{Q}(t) : \mathbb{Q}(u)] \leq 4$$

Now, fix  $u$  by  $C_2 \times C_2$

Hence;

$$\mathbb{Q}(u) \subset \mathbb{Q}(t)^{C_2 \times C_2}$$

Further by using the fixed field theorem;

$$[\mathbb{Q}(t) : \mathbb{Q}(t)^{C_2 \times C_2}] = 4$$

This follows that;

$$\mathbb{Q}(t) = \mathbb{Q}(t)^{C_2 \times C_2}$$

Claim: to find the subfield of  $\mathbb{Q}(u)$  that is fixed by all of  $A_4$  and since,  $\sigma$  and  $C_2 \times C_2$  generate  $A_4$ , then it remains to show that the subfield of  $\mathbb{Q}(u)$  is fixed by  $\sigma$ ;

$$\begin{aligned}\sigma(u) &= \sigma(t^2) \\ &= \left(\frac{t+i}{t-i}\right)^2 + \left(\frac{t-i}{t+i}\right)^2 \\ &= \frac{2(t^4 - 6t^2 + 1)}{(t^2 + 1)^2} \\ &= \frac{2(u-6)}{u+2}\end{aligned}$$

And;

$$\begin{aligned}\sigma^2(u) &= \sigma^2(t)^2 + \frac{1}{\sigma(t)^2} \\ &= \left(\frac{t+i}{t-i}\right)^2 + \left(\frac{t-i}{t+i}\right)^2 \\ &= \frac{-2(t^4 + 6t^2 + 1)}{(t^2 - 1)^2} \\ &= \frac{-2(u+6)}{u-2}\end{aligned}$$

Hence;

$$(x-u)(x-\sigma(u))(x-\sigma^2(u)) = x^3 - \frac{u(u^2-36)}{u^2-4}x^2 - 36x + \frac{4u(u^2-36)}{u^2-4}$$

Now, let;

$$v = \frac{u(u^2-36)}{u^2-4}$$

Here,  $v$  is fixed by  $\sigma$  and;

$$[\mathbb{Q}(u) : \mathbb{Q}(v)] \leq 3$$

Thus, by the fixed field theorem;

$$\begin{aligned}\mathbb{Q}(v) &= \mathbb{Q}(u)^{\langle \sigma \rangle} \\ &= \mathbb{Q}(t)^4\end{aligned}$$

Explicitly as;

$$v = \frac{(t^4+1)(t^8-34t^4+1)}{t^2(t^4-1)^2}$$

Hence, the required field is 
$$v = \frac{(t^4+1)(t^8-34t^4+1)}{t^2(t^4-1)^2}$$

### 3. a

Consider a field extension  $L/K$  which is algebraic if every element of  $L$  is algebraic over  $K$  that is every element of  $L$  is algebraic over  $K$ . Field extensions that are not algebraic which contain transcendental elements are called transcendental

Let  $F = \mathbb{C}(t)$  be the field of rational functions in  $t$

To prove: That every element of  $F$  that is not in  $\mathbb{C}$  is transcendental over  $\mathbb{C}$

Suppose that;

$$v \in \mathbb{C}(t) \setminus \mathbb{C}$$

This be the field of rational

Consider  $t$  is transcendental then there exists;

$$f(x), g(x) \in \mathbb{C}[x]$$

Such that;

$$g(t) \neq 0$$

And;

$$v = \frac{f(t)}{g(t)}$$

Since,  $v \notin \mathbb{C}$

The rational function  $\frac{f(x)}{g(x)}$  is non-constant

Now, on the contrary suppose that  $v$  is algebraic over  $\mathbb{C}$

Then there is a non-constant polynomial;

$$h(x) \in \mathbb{C}[x]$$

With degree  $n \geq 1$

And with;

$$h(v) = 0$$

Also;

$$v = \frac{f(t)}{g(t)}$$

Then following is obtained;

$$h\left(\frac{f(t)}{g(t)}\right) = 0$$

And, since;

$$g(x)^n \times h\left(\frac{f(x)}{g(x)}\right) \in \mathbb{C}[x]$$

This is a polynomial which is satisfied by  $t$

Hence,  $t$  must be algebraic over  $\mathbb{C}$

This is contradiction

**Therefore, every element of  $F$  that is not in  $\mathbb{C}$  is transcendental over  $\mathbb{C}$**

## Section 6

1. a

Consider the provided statement to check  $\sqrt{-31}$  is in the field  $\mathbb{Q}(\alpha)$  and is it in  $K$ .

Polynomial function is  $x^3 + x + 1$  and let  $\alpha$  be a complex root of the polynomial function. From the multiplicative property of degree  $\sqrt{-31} \notin \mathbb{Q}(\alpha)$  as  $\alpha$  has degree 3 but  $\sqrt{-31}$  has degree 2 over  $\mathbb{Q}$ .

[Comment](#)

Step 2 of 2 ^

If  $h(x) = x^3 + px + q$  is a polynomial function then their discriminants is provided as

$D = -4p^3 - 27q^2$ . Then in the polynomial function  $x^3 + x + 1$  where  $p = 1, q = 1$  so the discriminant will be as below:

$$\begin{aligned} D &= -4(1)^3 - 27(1)^2 \\ &= -4 - 27 \\ &= -31 \end{aligned}$$

However, as  $\sqrt{-31} \in K$  because the discriminant of  $x^3 + x + 1$  is  $-31$  as calculated above and the square root of the discriminant is a product of differences of elements in  $K$ .

2. a

To determine  $[K : \mathbb{Q}]$ , prove that  $K$  is a Galois extension of  $\mathbb{Q}$ , and determine its Galois group.

[Comment](#)

Step 2 of 2 ^

Let,

$$K = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$$

Then  $K$  is the splitting field of polynomial over  $\mathbb{Q}$  of  $(x^2 - 2)(x^2 - 3)(x^2 - 5)$ .

The Galois group is generated by  $\sigma$ ,  $\tau$  and  $\rho$  which are respectively interchange with the roots of  $x^2 - 2$ ,  $x^2 - 3$  and  $x^2 - 5$ .

Every element has exponent 2, then  $[K : \mathbb{Q}] = 2 \times 2 \times 2 = 8$ . Since the dimension of

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5}) \text{ as a } \mathbb{Q}\text{-vector space is } 8.$$

That is,

$$\{1, \sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{2}\sqrt{3}, \sqrt{2}\sqrt{5}, \sqrt{3}\sqrt{5}, \sqrt{2}\sqrt{3}\sqrt{5}\}$$

Thus,  $\{1, \sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{2}\sqrt{3}, \sqrt{2}\sqrt{5}, \sqrt{3}\sqrt{5}, \sqrt{2}\sqrt{3}\sqrt{5}\}$  is a  $\mathbb{Q}$ -basis for  $K$ .

Therefore,  $K$ -automorphism extension of  $\mathbb{Q}$  is  $[K : \mathbb{Q}] = 8$ . So if the Galois group is

$$G = \text{Gal}(K / \mathbb{Q}) \text{ then } |G| = 8.$$

3. a

Field is defined as the nonzero commutative division ring or can be defined as the ring whose nonzero element forms an abelian group.

[Comment](#)

Step 2 of 3 ^

Consider  $K \supset L \supset F$  be a chain of extension fields of degree 2.

To show: that  $K$  can be generated over  $F$  by the root of an irreducible quartic polynomial of the form  $x^4 + bx^2 + c$

For the proof consider  $L/F$  to be extension of degree 4.

Then;

$$L = F(\alpha)$$

Where,  $\alpha$  is a root of a polynomial of the form  $x^4 + bx^2 + c$

Now, if;

$$\alpha^2 \in F$$

Then;

$$f(x) = (x - \alpha^2)(x - \beta^2)$$

This does not factorizes in  $K$

Thus, the polynomial is irreducible

$$\text{If } L = F(\alpha)$$

Then;

$$\alpha = \pm \frac{-a \pm \sqrt{a^2 - 4b}}{2}$$

And, hence;

$$F(\alpha^2) = F(\sqrt{a^2 - 4b})$$

This should be a quadratic extension of  $F$

That is  $F(\alpha)$  would be of degree 2 over  $F$  and not 4.

And, if  $L/F$  is an extension of degree 4 having a quadratic intermediate field;

$$K = F(\sqrt{d})$$

Then,  $L$  is a quadratic extension of  $K$  so,

$$L = K(\sqrt{\beta}), \beta \in K$$

That is;

$$\begin{aligned} \beta &= r + s\sqrt{d} \\ &\in K \end{aligned}$$

Thus,  $\alpha$  is the root

**Therefore,  $K$  can be generated over  $F$  by the root of an irreducible quartic polynomial of the form  $x^4 + bx^2 + c$**

## Section 7

1. a



**Solution:** We will determine the intermediate fields of an extension field of the form  $F(\sqrt{a}, \sqrt{b})$ .

Let us assume that  $\sqrt{b} \notin F(\sqrt{a})$  and  $\sqrt{a} \notin F(\sqrt{b})$ .

Then notice that it is obvious

$$[F(\sqrt{a}, \sqrt{b}) : F] = 4.$$

Therefore any intermediate fields must be quadratic extensions of  $F$ . Therefore notice that  $F(\sqrt{a})$ ,  $F(\sqrt{b})$ , and  $F(\sqrt{a}, \sqrt{b})$  are three such intermediate fields explained above.

Now if there is any other intermediate field besides  $F(\sqrt{a})$ ,  $F(\sqrt{b})$ , and  $F(\sqrt{a}, \sqrt{b})$  it must be generated by an element of the form

$$\chi = \alpha\sqrt{a} + \beta\sqrt{b} + \delta\sqrt{ab}, \text{ where } \alpha, \beta, \delta \in F.$$

Now we have

$$\begin{aligned} \chi^2 &= (\alpha\sqrt{a} + \beta\sqrt{b} + \delta\sqrt{ab})^2 \\ &= (\alpha^2a + \beta^2b + \delta^2ab) + 2\beta\delta b\sqrt{a} + 2\alpha\delta a\sqrt{b} + 2\alpha\beta\sqrt{ab}. \end{aligned}$$

Now if we have

$$[\alpha : \beta : \delta] = [2\beta\delta b : 2\alpha\delta a : 2\alpha\beta]$$

then  $\chi$  will have degree 2.

Now notice that if  $\alpha\beta\delta = 0$ , then it is easy to see that this condition forces at least one other to vanish, so one of the three intermediate fields  $F(\sqrt{a})$ ,  $F(\sqrt{b})$ , and  $F(\sqrt{a}, \sqrt{b})$  is obtained.

Otherwise let us assume

$$\alpha^2\delta a = \beta^2\delta b.$$

Then we have

$$\alpha\sqrt{a} = \pm\beta\sqrt{b}.$$

Now notice that if  $\alpha\beta \neq 0$ , then we contradict the fact that  $\sqrt{a} \notin F(\sqrt{b})$ .

Therefore there are no other intermediate fields other than  $F(\sqrt{a})$ ,  $F(\sqrt{b})$ , and  $F(\sqrt{a}, \sqrt{b})$ .

This completes the solution.

## Result

2 of 2

The intermediate fields of an extension field of the form  $F(\sqrt{a}, \sqrt{b})$  are  $F(\sqrt{a})$ ,  $F(\sqrt{b})$ , and  $F(\sqrt{a}, \sqrt{b})$ .

## 2. a

Fundamental theorem on Galois Theory states that given a field extension  $E/F$  that is finite and Galois there is one-one correspondence between its intermediate fields and subgroups of its Galois group

Let  $K/F$  be a Galois extension such that;

$$G(K/F) \approx C_2 \times C_{12}$$

a.

To find: The number of intermediate field of  $L$  where  $[L:F] = 4$

Since, given that;

$$G(K/F) \approx C_2 \times C_{12}$$

Therefore, number of elements will be;

$$\begin{aligned} C_2 \times C_{12} &= 2 \times 12 \\ &= 24 \end{aligned}$$

And also;

$$[L:F] = 4$$

Now, the number of intermediate field will be the number of subgroups having the number elements;

$$\begin{aligned} \frac{C_2 \times C_{12}}{[L:F]} &= \frac{24}{4} \\ &= 6 \end{aligned}$$

Now, the field's having six elements are;

$$\langle (0,2) \rangle, \langle (1,2) \rangle \text{ and } \mathbb{Z}_2 \times \langle 4 \rangle$$

**Therefore, there are three intermediate field of  $L$  where  $[L:F] = 4$**

b.

To find: The number of intermediate field of  $L$  where  $[L:F] = 9$

Since, given that;

$$G(K/F) \approx C_2 \times C_{12}$$

Therefore, number of elements will be;

$$\begin{aligned} C_2 \times C_{12} &= 2 \times 12 \\ &= 24 \end{aligned}$$

And also;

$$[L:F] = 9$$

Here, clearly  $9 \nmid 24$

**Therefore, there is no field with  $[L:F] = 9$**

c.

To find: The number of intermediate field of  $L$  where  $Gal(L/M) \approx C_4$

Since, given that;

$$G(K/F) \approx C_2 \times C_{12}$$

Therefore, number of elements will be;

$$\begin{aligned} C_2 \times C_{12} &= 2 \times 12 \\ &= 24 \end{aligned}$$

Now, it is required to find the number of subgroups of  $C_2 \times C_{12}$ , that is isomorphic to  $C_4$

That is;

$$\langle (0,2) \rangle \text{ and } \langle (1,2) \rangle$$

These are the subgroups of  $C_2 \times C_{12}$ , that is isomorphic to  $C_4$

**Therefore, there are two intermediate field of  $L$  where  $Gal(L/M) \approx C_4$**

3. a

If  $L$  is an extension of  $F$  which is in turn an extension of  $K$ , then  $F$  is said to be an intermediate field of the field extension  $L/K$

a.

To find: the number of intermediate fields  $L$  with  $[L:F] = 2$  will be there when  $K/F$  is a Galois extension with Alternating group  $A_4$

Here, use the result that, an extension  $K/F$  such that  $[K:F] = n$  and a positive integer  $d$  dividing  $n$  such that there is no intermediate field  $K \subset L \subset K$  with  $[L:F] = d$

The group  $A_4$  is defined as the group of even permutation on four elements.

Now,  $A_n$  is defined as;

$$A_n = \frac{n!}{2}$$

Here,  $n$  is the order of  $A_4$

Since, the cardinality of  $A_4$  is 4

Thus,  $n = 4$

That is;

$$\begin{aligned} |A_4| &= \frac{4!}{2} \\ &= \frac{1 \times 2 \times 3 \times 4}{2} \\ &= 12 \end{aligned}$$

Since,  $[L:F] = 2$

So, by using the above stated result the number of intermediate fields will be;

$$\frac{12}{2} = 6$$

**Therefore, the number of intermediate fields  $L$  with  $[L:F] = 2$  will be there when,  $K/F$  is a Galois extension with Alternating group  $A_4$  is  $\boxed{6}$**

b.

To find: the number of intermediate fields  $L$  with  $[L:F] = 2$  will be there when,  $K/F$  is a Galois extension with Dihedral group  $D_4$

Here, use the result that, an extension  $K/F$  such that  $[K:F] = n$  and a positive integer  $d$  dividing  $n$  such that there is no intermediate field  $K \subset L \subset K$  with  $[L:F] = d$

The Dihedral group  $D_4$  is one of the two non-Abelian groups of the five groups which is total of group order 8

That is,

$$|D_4| = 8$$

Since,  $[L:F] = 2$

So, by using the above stated result the number of intermediate fields will be;

$$\frac{8}{2} = 4$$

**Therefore, the number of intermediate fields  $L$  with  $[L:F] = 2$  will be there when,  $K/F$  is a Galois extension with Dihedral group  $D_4$  is  $\boxed{4}$**

4. a

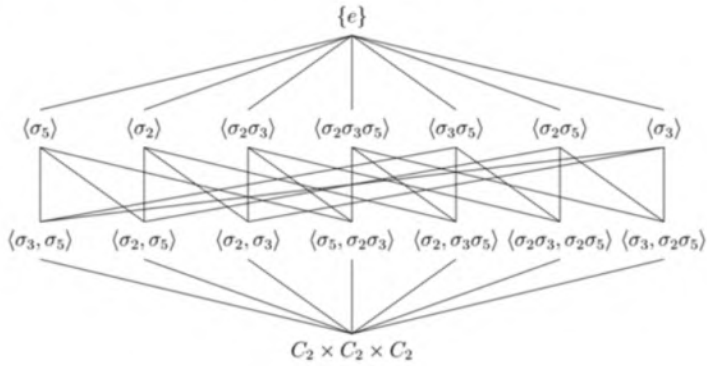
Consider the provided statement to determine all intermediate fields of the given expression. It is provided that  $F = \mathbb{Q}$  and  $K = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$ .

It is assumed that  $K$  is the splitting field of the polynomial  $(x^2 - 2)(x^3 - 3)(x^2 - 5)$  and  $F \subset K$  is a Galois extension of order 8 where the Galois group is provided as below:

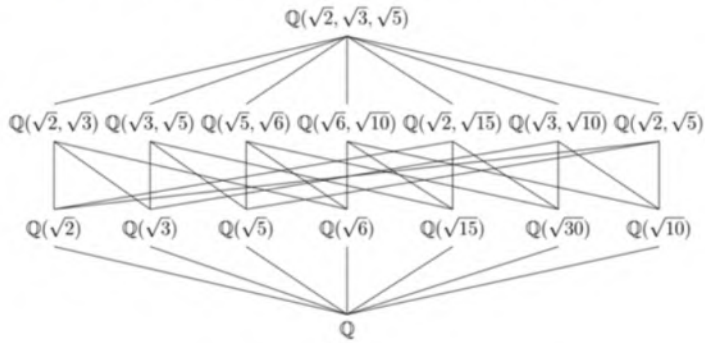
$$G = C_2 \times C_2 \times C_2 \\ = \langle \sigma_2, \sigma_3, \sigma_5 \rangle$$

Where  $\sigma_k$  is the field automorphism over  $\mathbb{Q}$  such that the map under consideration sends the element  $\sqrt{k}$  to the element  $-\sqrt{k}$ .

The lattice diagram for the subgroup of  $G = C_2 \times C_2 \times C_2$  is given as below:



It corresponds to the lattice diagram of intermediate fields when it is computing as fixed fields according to the theorem 16.7.1, then lattice diagram is provided as below:



5. a

Cubic irreducible polynomial is defined as, if  $F$  is a field, a non-constant polynomial is irreducible over  $F$  if its coefficients belong to  $F$  and it cannot be factored into the product of two non-constant polynomials with coefficients in  $F$ .

[Comment](#)

Step 2 of 3 ^

Let  $f(x)$  be an irreducible cubic polynomial over  $\mathbb{Q}$  whose Galois group is  $S_3$

To determine: the possible Galois groups of the polynomial  $(x^3 - 1)f(x)$

For the proof consider  $F$  be the splitting field of  $f$  over  $\mathbb{Q}$  and  $E$  the splitting field of  $x^3 - 1$

This means that  $E$  is the splitting field of  $x^2 + x + 1$

The splitting field of  $(x^3 - 1)f(x)$  will be the smallest extension containing both the roots of  $f$  and those of  $x^2 + x + 1$

Let;

$$G = G(\text{Smallest extension})/\mathbb{Q}$$

Now, for the case if;

$$f = x^3 - 2$$

Then;

$$G = S_3$$

Else,  $E \cap F$  has dimension 1 over  $\mathbb{Q}$

That is;

$$E \cap F = \mathbb{Q}$$

And, thus;

$$\begin{aligned} G &= G(E/\mathbb{Q}) \times G(F/\mathbb{Q}) \\ &= (\mathbb{Z}/2\mathbb{Z}) \times S_3 \end{aligned}$$

Thus, the possible Galois groups of the polynomial  $(x^3 - 1)f(x)$  is  $(\mathbb{Z}/2\mathbb{Z}) \times S_3$

6. a

**Given:**  $K/F$  is a Galois extension whose Galois group is the symmetric group  $S_3$ .

**Solution:** We will propose to prove that  $K$  is splitting field of an irreducible cubic polynomial over  $F$ .

Notice that  $S_3$  is a group of order 6 and consider the subgroup generated by the element  $(1\ 2)$  in  $S_3$ . This proves the existence of index 3 subgroup of  $S_3$ .

Then there is an intermediate field  $L$  with degree 3 over  $F$ .

Now notice that since 3 is prime, we have

$$L = F(\alpha), \quad \text{where } \alpha \in L - F.$$

Let us now assume that  $f$  be the minimal polynomial of  $\alpha$  over  $F$ . Then  $f$  is a cubic polynomial.

Now recall that  $S_3$  has only one non-trivial proper normal subgroup which is  $A_3$  of order 3. So  $S_3$  has no normal subgroup of index 3.

This follows that  $L/F$  is not a Galois extension.

Hence by the definition  $f$  does not split completely in  $L$ .

Now note that  $K$  is a splitting field over  $F$  and  $f$  has a root in  $K$ , it follows that  $f$  splits completely in  $K$ .

Now notice that the field generated over  $F$  by the roots of  $f$  is strictly larger than  $L$ .

Now we have

$$[K : F] = 2.$$

Therefore his field can only be  $K$ . Therefore  $K$  is the splitting field of the irreducible cubic  $f$ .

This completes the proof.

## Result

2 of 2

Considering the note that  $S_3$  has no normal subgroup of index 3, we have assert that  $K$  is a splitting field of an irreducible cubic polynomial.

7. a

**Solution:**

(a) We will determine the irreducible polynomial for  $i + \sqrt{2}$  over  $\mathbb{Q}$ . Now notice that

$$\mathbb{Q} \subseteq \mathbb{Q}(i + \sqrt{2}) \quad \text{and} \quad \mathbb{Q}(i + \sqrt{2}) \subseteq \mathbb{Q}(i, \sqrt{2}).$$

Therefore we have

$$\mathbb{Q} \subseteq \mathbb{Q}(i + \sqrt{2}) \subseteq \mathbb{Q}(i, \sqrt{2}). \quad (1)$$

Now we will prove that  $[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}] = 4$ .

**Claim:** Let  $F$  be a splitting field for  $f(x) = x^4 + 1$ . Then we have order of the group  $G_{F/\mathbb{Q}}$  is 4.



**Proof of the Claim:** Notice that the polynomial  $f(x)$  has roots  $\{\sqrt[4]{-1}, \chi_4 \sqrt[4]{-1}, \chi_4^2 \sqrt[4]{-1}, \chi_4^3 \sqrt[4]{-1}\}$ .

Now notice that

$$\begin{aligned}\sqrt[4]{-1} &= i^{\frac{1}{2}} \\ &= \exp\left(\frac{i\pi}{4}\right) \\ &= \frac{\sqrt{2} + i\sqrt{2}}{2} \\ &= \frac{(1+i)\sqrt{2}}{2}.\end{aligned}$$

Since  $\chi_4 = i$ , a splitting field for  $f(x)$  is given by  $F = \mathbb{Q}(i, \sqrt{2})$ . Now observe that

$$[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}] = [\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \times 2 = 4.$$

Since we know that  $[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}(\sqrt{2})] = 2$  and  $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ . Hence order of the group  $G_{F/\mathbb{Q}}$  is 4.

We proved the Claim.

Now from the above Claim we observe that  $[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}] = 4$ .

Therefore from (1) it follows that the degree of  $\mathbb{Q}(i + \sqrt{2})$  is either 2 or 4.

Now note that  $i + \sqrt{2}$  can not have degree 2 over  $\mathbb{Q}$ . Hence the degree of  $\mathbb{Q}(i + \sqrt{2})$  is 4 over  $\mathbb{Q}$ .

The minimal polynomial is therefore given by

$$(x - i - \sqrt{2})(x - i + \sqrt{2})(x + i - \sqrt{2})(x + i + \sqrt{2}) = x^4 - 2x^2 + 9.$$

(b) We will propose to prove that the set  $\{1, i, \sqrt{2}, i\sqrt{2}\}$  is a basis of  $\mathbb{Q}(i, \sqrt{2})$  over  $\mathbb{Q}$ .

Note that by the above claim we have

$$[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}] = 4.$$

Now trivially every element of  $\mathbb{Q}(i, \sqrt{2})$  can be written as a linear combination of the elements of the set  $\{1, i, \sqrt{2}, i\sqrt{2}\}$ .

Therefore  $\{1, i, \sqrt{2}, i\sqrt{2}\}$  forms the minimal spanning set for  $\mathbb{Q}(i, \sqrt{2})$  over  $\mathbb{Q}$ .

Hence the set  $\{1, i, \sqrt{2}, i\sqrt{2}\}$  is a basis of  $\mathbb{Q}(i, \sqrt{2})$  over  $\mathbb{Q}$ .

This completes the proof.

## Result

3 of 3

First we prove that  $x^4 - 2x^2 + 9$  is the required minimal polynomial and then proved that the set  $\{1, i, \sqrt{2}, i\sqrt{2}\}$  is a basis of  $\mathbb{Q}(i, \sqrt{2})$  over  $\mathbb{Q}$ .

8. a

Galois group is defined as the group of the field Automorphism under the given criteria of the composition.

[Comment](#)

Step 2 of 5 ^

To determine: the irreducible factors over each of the fields

The splitting field of  $x^4 + 2$  is;

$$\mathbb{Q}(\alpha, i)$$

Now, the corresponding extension is given as;

$$\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}(\sqrt{2})$$

This extension has order 4

Now, the automorphisms of the Galois group are determined by;

$$\sqrt[4]{2} \rightarrow \pm \sqrt[4]{2}$$

And;

$$i \rightarrow \pm i$$

Therefore, the required Galois group is  $\boxed{\mathbb{Z}_2 \times \mathbb{Z}_2}$

The splitting field of  $x^4 + 2$  is;

$$\mathbb{Q}(\alpha, i)$$

Now, the corresponding extension is given as;

$$\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}(\sqrt{2}, i)$$

This extension has order 2

Now, the automorphisms of the Galois group are determined by;

$$\sqrt[4]{2} \rightarrow \pm \sqrt[4]{2}$$

And;

$$i \rightarrow \pm i$$

Therefore, the required Galois group is  $\boxed{\mathbb{Z}_2}$

The splitting field of  $x^4 + 2$  is;

$$\mathbb{Q}(\alpha, i)$$

Now, the corresponding extension is given as;

$$\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}(\alpha)$$

This extension has order 2

Now, the automorphisms of the Galois group are determined by;

$$\sqrt[4]{2} \rightarrow \pm \sqrt[4]{2}$$

And;

$$i \rightarrow \pm i$$

Therefore, the required Galois group is  $\boxed{\mathbb{Z}_2}$

The splitting field of  $x^4 + 2$  is;

$$\mathbb{Q}(\alpha, i)$$

Now, the corresponding extension is given as;

$$\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}(\alpha, i)$$

This extension is trivial

Therefore, the required Galois group is  $\langle e \rangle$

9. a

A Galois group is defined as any group that forms an field automorphisms under the given composition

[Comment](#)

Step 2 of 4 ^

Let  $\zeta = e^{2\pi i/5}$

To prove: That  $K = \mathbb{Q}(\zeta)$  is a splitting field for the polynomial  $x^5 - 1$  over  $\mathbb{Q}$  and to determine the degree of  $[K : \mathbb{Q}]$

Since;

$$\zeta^5 - 1 = 0$$

That is;

$$\begin{aligned} x^5 - 1 &= (x-1)(x^4 + x^3 + x^2 + x + 1) \\ &= (x-1)(x-\zeta)(x-\zeta^2)(x-\zeta^3)(x-\zeta^4) \end{aligned}$$

Thus,  $K$  is  $x^5 - 1$

**Therefore,  $K = \mathbb{Q}(\zeta)$  is a splitting field for the polynomial  $x^5 - 1$  over  $\mathbb{Q}$**

That is this is a splitting field and since  $\zeta \neq 1$ ,  $\zeta$  is the root of an irreducible that is a cyclotomic polynomial, this polynomial is of degree 4

Thus,  $[K : \mathbb{Q}] = 4$

To prove: that  $K$  is a Galois extension of  $\mathbb{Q}$  and to determine its Galois group

Let  $\sigma_i$  send  $\zeta$  to  $\zeta^i$

Then  $\sigma_1, \dots, \sigma_4$  are all automorphisms of  $K$ .

Since,

$$\begin{aligned} |\text{Aut}(K/\mathbb{Q})| &= 4 \\ &= [K:\mathbb{Q}] \end{aligned}$$

Now,  $K/\mathbb{Q}$  is Galois, and the Galois group is  $\mathbb{Z}/4\mathbb{Z}$ .

Therefore,  $K$  is a Galois extension of  $\mathbb{Q}$ .

[Comment](#)

Step 4 of 4 ^

Also;

$$\begin{aligned} \sigma_2^4 &= \sigma_1(16) \\ &= \sigma_3(8)\sigma_2 \\ &= \sigma_4^2 \end{aligned}$$

Thus,  $\text{Gal}(K/\mathbb{Q})$  is of order 4 and has an element of order 4

Thus, it cannot be  $V_4$  and must be  $\mathbb{Z}/4\mathbb{Z}$ .

Hence, its Galois group is  $\mathbb{Z}/4\mathbb{Z}$ .

10. a

Consider the provided statement to prove the given condition.

As it is provided that  $K/F$  be a Galois extension with the Galois group and also  $H$  be a subgroup of  $G$ .

It is assumed that  $K, F$  have characteristic zero and  $K/F$  is a finite extension. Therefore, the chain of extensions is considered as  $F \subset K'' \subset K$ . From the primitive element theorem there exists  $\beta \in K''$  such that  $K'' = F(\beta)$ .

[Comment](#)

Step 2 of 2 ^

It is claimed that  $G_\beta = H$  therefore clearly  $H \subset G_\beta$  so it is supposed that  $\sigma \in \frac{G_\beta}{H}$ .

Then from Corollary 16.7.2,  $H' := H \langle \sigma \rangle$  and  $K''' \subset K''$ .

Every element in  $K''$  is fixed by  $H'$  so  $K''' = K''$  but then from Corollary 16.7.2,

$$\begin{aligned} [K:K''] &= |H| \\ &= |H'| \end{aligned}$$

Therefore, it contradicts that  $H \subsetneq H'$  and hence given statement is **proved**.

11. a

Splitting field of a polynomial with coefficients in a field is defined as the smallest field extension of that field over which the polynomial decomposes into linear factors.

Let  $\alpha = \sqrt[3]{2}$ ,  $\beta = \sqrt{3}$ , and  $\gamma = \alpha + \beta$ . Let  $L$  be the field  $\mathbb{Q}(\alpha, \beta)$ , and let  $K$  be the splitting field of the polynomial  $(x^3 - 2)(x^3 - 3)$  over  $\mathbb{Q}$ .

a.

To determine: the irreducible polynomial  $f$  for  $\gamma$  over  $\mathbb{Q}$  and its roots in  $\mathbb{C}$ .

For the proof first consider;

$$\begin{aligned} (\gamma - \beta)^3 &= \gamma^3 - 3\beta\gamma^2 + 9\gamma - 3\beta \\ &= 2 \end{aligned}$$

Hence;

$$\beta = \frac{\gamma^3 + 9\gamma - 2}{3\gamma^2 + 3}$$

So;

$$\mathbb{Q}(\gamma) = \mathbb{Q}(\alpha, \beta)$$

Since,  $\alpha$  has degree 3 and  $\beta$  has degree 2 over  $\mathbb{Q}$ , then by using the result of the corollary which states that;

Let  $K$  be an extension field of a field  $F$ , let  $K$  and  $F'$  be subfields of  $K$  that are finite extensions of  $F$ , and let  $K'$  denote the subfield of  $K$  generated by the two fields  $K$  and  $F'$  together. Let;

$$[K' : F] = N$$

$$[K : F] = m$$

$$[F' : F] = n$$

Then,  $m$  and  $n$  divide  $N$  and  $N \leq mn$

So, by using the above result;

$$[L : \mathbb{Q}] = 6$$

Thus, the irreducible polynomial for  $\gamma$  has degree 6. Now, let;

$$f(x) = ((x - \beta)^3 - 2)((x + \beta)^3 - 2)$$

$$f(\gamma) = 0$$

By construction;

$$\begin{aligned} f(x) &= (x^3 + 9x - 2 - 3\beta(x^2 + 1))(x^3 + 9x - 2 + 3\beta(x^2 + 1)) \\ &= (x^3 + 9x - 2)^2 - 27(x^2 + 1)^2 \\ &= x^6 - 9x^4 - 4x^3 + 27x^2 - 36x - 23 \end{aligned}$$

Since, degree of  $f(x) = 6$

Thus, it is an irreducible polynomial for  $\gamma$

Now,  $z \in \mathbb{C}$  is a root of  $f(x)$  if and only if;

$$z \pm \beta = \omega^j \alpha$$

Where,

$$\omega = e^{2\pi i/3}$$

$$j \in \{0, 1, 2\}$$

Hence, the roots of  $f(x)$  are  $\pm\beta + \omega^j \alpha$

b.

To determine: the Galois group of  $K/\mathbb{Q}$

For the proof first denote;

$$G = G(K/\mathbb{Q})$$

So, from the first part,  $f(x)$  splits completely in  $K$  by the splitting theorem

Similarly,  $(x^3 - 2)(x^3 - 3)$  splits completely in the splitting fields for  $f(x)$  since  $\alpha, \beta$  can be obtained as linear combinations of the roots of  $f(x)$

Thus,  $K$  is the splitting field for  $f(x)$

If  $j \in \{1, 2\}$ ;

$$\begin{aligned} \pm\beta + \omega^j \alpha &= \pm\beta + \frac{-1 - (-1)^j i\beta}{2} \alpha \\ &= \pm\beta - \frac{1}{2} \alpha - \frac{(-1)^j}{2} i\beta \end{aligned}$$

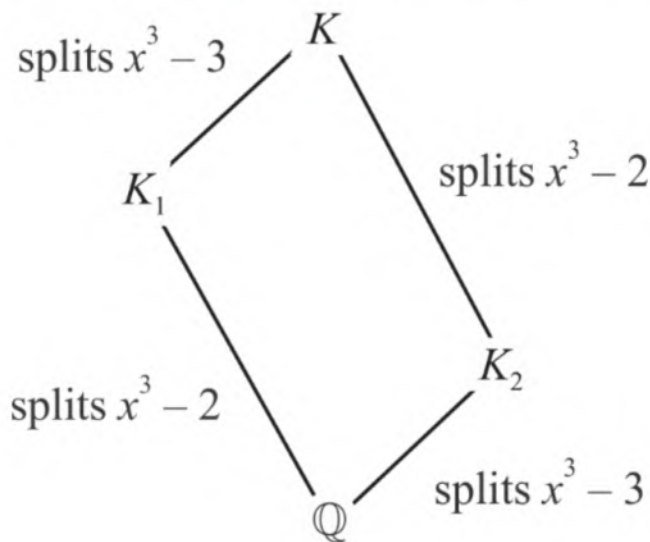
Hence;

$$\begin{aligned} K &= \mathbb{Q}(\gamma, i) \\ &= \mathbb{Q}(\alpha, \beta, i) \end{aligned}$$

And so by using the multiplicative property of the degree;

$$\begin{aligned} [K : \mathbb{Q}] &= [K : \mathbb{Q}(\gamma)][\mathbb{Q}(\gamma) : \mathbb{Q}] \\ &= 12 \end{aligned}$$

Now, let  $K_1$  be the splitting field for  $x^3 - 2$  and let  $K_2$  be the splitting field for  $x^3 - 3$ . So, the following lattice will be formed;



Further by using the theorem which states that;

Let  $K/F$  be a finite extension and let  $G$  be its Galois group. Then  $K$  is a splitting field over  $F$

So, by using this result; the subgroups;

$$\begin{aligned} G_1 &= G(K/K_1) \\ G_2 &= G(K/K_2) \end{aligned}$$

In fixing these subgroups of  $G$  by  $K_1, K_2$  respectively are normal subgroups of  $G$ . This is true by using the result of the theorem which states that;

Let  $K/F$  be a Galois extension with Galois group  $G$ , and let  $L$  be the fixed field  $L$  of a subgroup  $H$  of  $G$ . The extension  $L/F$  is a Galois extension if and only if  $H$  is a normal subgroup of  $G$ . If so, the Galois group  $G(L/F)$  is isomorphic to the quotient group  $G/H$



Claim:

$$\begin{aligned} G &\approx G_1 \times G_2 \\ &\approx C_2 \times S_3 \end{aligned}$$

For this consider the multiplication map;

$$\mu: G_1 \times G_2 \rightarrow G$$

First claim that  $\mu$  is injective

So, by using the result of the theorem which states that;

Let  $K/F$  be a Galois extension with Galois group  $G$ , and let  $g$  be a polynomial with coefficients in  $F$  that splits completely in  $K$ . Let its roots in  $K$  be  $\beta_1, \dots, \beta_r$  then, if  $K$  is a splitting field of  $g$  over  $F$ , the operation on the roots is faithful and by its operation on the roots,  $G$  embeds as a subgroup of symmetric group  $S_r$ .

Now, it remains to show that any  $\sigma \in G_1 \cap G_2$  operates as the identity on  $\pm\beta + \omega^j\alpha$

Since,

$$\begin{aligned} \omega^j\alpha &\in K_1 \\ \pm\beta &\in K_2 \end{aligned}$$

Hence;

$$G_1 \cap G_2 = \{1\}$$

Now, also,

$$\begin{aligned} |G_1 \times G_2| &= 12 \\ &= |G| \end{aligned}$$

Thus,  $\mu$  is also surjective, and is therefore an isomorphism

## Section 8

### 1. a

The Klein four groups are defined as the direct product of the group with itself and this group consists of the elements under coordinate wise-multiplication.

[Comment](#)

Step 2 of 3 ^

Let  $K/F$  be a Galois extension whose group  $G$  is a Klein four group  $D_2$

To prove: that  $K$  can be obtained by adjoining two square roots to  $F$  and how  $G$  acts on  $K$

For the proof use the result of Main theorem which states that;

Let  $K$  be a Galois extension of a field  $F$  and let  $G$  be its Galois group. there is a bijective correspondence between subgroups of  $G$  and intermediate fields and this correspondence associates to a subgroup  $H$  its fixed field, and to an intermediate field  $L$ , the Galois group of  $K$  over  $L$ . Then the maps;

$$\begin{aligned} H &\rightarrow K^H \\ L &\rightarrow G(K/L) \end{aligned}$$

These are inverse functions.

Thus by using the result of above defined main theorem, if the Galois group has three subgroups of index 2 then  $K$  contains three subfields containing  $F$  which have degree 2 over  $F$

Further, let two of these subfields be  $F(\beta)$  and  $F(\gamma)$

Since, these subfields are distinct and both have degree 2 over  $F$ ,  $\beta \notin F(\gamma)$  and  $\gamma \notin F(\beta)$  but  $\beta, \gamma \in K$

Thus,  $F(\beta, \gamma)$  has degree 2 over  $F(\beta)$ ,  $F(\beta, \gamma)$  has degree 4 over  $F$  and  $K \supset F(\beta, \gamma)$

This implies that;

$$K = F(\beta, \gamma)$$

Thus,  $K$  is the splitting field for;

$$(x - \beta)(x + \beta)(x - \gamma)(x + \gamma)$$

Where, splits in  $F(\beta)$  and  $(x - \gamma)(x + \gamma)$  splits in  $F(\gamma)$

Thus,  $K/F$  is a biquadratic extension

**Therefore,  $K$  can be obtained by adjoining two square roots to  $F$  and there is a bijective correspondence between subgroups of  $G$  and intermediate field  $K$**

## 2. a

A Galois group is defined as a field which will be obtained by field automorphisms and under the composition.

To determine: the Galois group of each of the following expression over  $\mathbb{Q}$

a.

Consider the expression;

$$x^3 - 2$$

Now, the roots that is the factors of the given polynomial are one real root that is,  $\sqrt[3]{2}$  and the two complex roots that is;  $-\sqrt[3]{2}, (-1)^{2/3} \sqrt[3]{2}$

Clearly, the factor does not lie in  $\mathbb{Q}$

Thus, the function is irreducible

Now, finding the determinant to look at if it is coming to be a square or not.

Now, the equations of the form;

$$f(x) = x^3 + q$$

The discriminant will be;

$$\Delta = -27q^2$$

Comparing the function  $x^3 - 2$ , then;

$$q = -2$$

Substituting the value;

$$\begin{aligned} \Delta &= -27q^2 \\ &= -27(-2)^2 \\ &= -108 \end{aligned}$$

This is not a square

Since, the function is irreducible and the determinant is also not a square.

**Therefore, the Galois group is  $S_4$**

b.

Consider the expression;

$$x^3 + 3x + 14$$

Now, the roots that is the factors of the given polynomial are one real root that is,  $-2$  and the two complex roots that is;  $1 - i\sqrt{6}, 1 + i\sqrt{6}$

Clearly, the factor does not lie in  $\mathbb{Q}$

Thus, the function is irreducible

Now, finding the determinant to look at if it is coming to be a square or not.

Now, the equations of the form;

$$f(x) = x^3 + px + q$$

The discriminant will be;

$$\Delta = -4p^3 - 27q^2$$

Comparing the function  $x^3 + 3x + 14$ , then;

$$p = 3$$

$$q = 14$$

Substituting the value;

$$\begin{aligned}\Delta &= -4p^3 - 27q^2 \\ &= -4(3)^3 - 27(14)^2 \\ &= -5400\end{aligned}$$

This is not a square

Since, the function is irreducible and the determinant is also not a square.

Therefore, the Galois group is  $S_4$

c.

Consider the expression;

$$x^3 - 3x^2 + 1$$

Now, the roots that is the factors of the given polynomial are three real root that is,  $-0.53209, 0.65270, 2.8794$

Clearly, the factor lies in  $\mathbb{Q}$

Thus, the function is reducible

Now, finding the determinant to look at if it is coming to be a square or not.

Now, the equations of the form;

$$f(x) = ax^3 + bx^2 + d$$

The discriminant will be;

$$\Delta = -4b^3d - 27a^2d^2$$

Comparing the function  $x^3 - 3x^2 + 1$ , then;

$$a = 1$$

$$b = -3$$

$$d = 1$$

Substituting the value;

$$\begin{aligned}\Delta &= -4b^3d - 27a^2d^2 \\ &= -4(-3)(1) - 27(1)^2(1)^2 \\ &= 12 - 27 \\ &= -15\end{aligned}$$

This is not a square

Since, the function is reducible and the determinant is also not a square.

Therefore, the Galois group is  $D_4$  or  $C_4$

d.

Consider the expression;

$$x^3 - 21x + 7$$

Now, the roots that is the factors of the given polynomial are one real root that is;

$$-4.7409, 0.33513, 4.4058$$

Clearly, the factor lies in  $\mathbb{Q}$

Thus, the function is reducible

Now, finding the determinant to look at if it is coming to be a square or not.

Now, the equations of the form;

$$f(x) = x^3 + px + q$$

The discriminant will be;

$$\Delta = -4p^3 - 27q^2$$

Comparing the function  $x^3 + 3x + 14$ , then;

$$p = -21$$

$$q = 7$$

Substituting the value;

$$\begin{aligned}\Delta &= -4p^3 - 27q^2 \\ &= -4(-21)^3 - 27(7)^2 \\ &= 35721\end{aligned}$$

This is a square

Since, the function is reducible and the determinant is a square.

**Therefore, the Galois group is  $D_2$**

e.

Consider the expression;

$$x^3 + x^2 - 2x - 1$$

Now, the roots that is the factors of the given polynomial are one real root that is;

$$-1.8019, -0.44504, 1.2470$$

Clearly, the factor lies in  $\mathbb{Q}$

Thus, the function is reducible

Now, finding the determinant to look at if it is coming to be a square or not.

Now, the equations of the form;

$$f(x) = ax^3 + bx^2 + cx + d$$

The discriminant will be;

$$\Delta = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd$$

Comparing the function  $x^3 + x^2 - 2x - 1$ , then;

$$a = 1$$

$$b = 1$$

$$c = -2$$

$$d = -1$$

Substituting the value;

$$\begin{aligned}\Delta &= b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd \\ &= (1)^2(-2)^2 - 4(1)(-2)^3 - 4(1)^3(-1) - 27(1)^2(-1)^2 + 18(1)(1)(-2)(-1) \\ &= 1\end{aligned}$$

This is a square

Since, the function is reducible and the determinant is a square.

**Therefore, the Galois group is  $D_2$**

f.

Consider the expression;

$$x^3 + x^2 - 2x + 1$$

Now, the roots that is the factors of the given polynomial are one real root that is  $-2.1479$ ; and two complex root  $0.57395 - 0.36899i, 0.57395 + 0.36899i$

Clearly, the factor does not lie in  $\mathbb{Q}$

Thus, the function is irreducible

Now, finding the determinant to look at if it is coming to be a square or not.

Now, the equations of the form;

$$f(x) = ax^3 + bx^2 + cx + d$$

The discriminant will be;

$$\Delta = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd$$

Comparing the function  $x^3 + x^2 - 2x + 1$ , then;

$$a = 1$$

$$b = 1$$

$$c = -2$$

$$d = 1$$

Substituting the value;

$$\begin{aligned}\Delta &= b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd \\ &= (1)^2(-2)^2 - 4(1)(-2)^3 - 4(1)^3(1) - 27(1)^2(1)^2 + 18(1)(1)(-2)(1) \\ &= -79\end{aligned}$$

This is not a square

Since, the function is irreducible and the determinant is also not a square.

Therefore, the Galois group is  $S_4$

3. a

A quadratic polynomial in  $x$  is defined as an expression of the form;

$$ax^2 + bx + c = 0$$

Where,  $a, b, c$  are constants

[Comment](#)

Step 2 of 3 ^

To determine: the quadratic polynomial  $q(x)$  that appears in  $f(x) = (x - \alpha_1)g(x)$  explicitly in terms of  $\alpha_1$  and the coefficients of  $f$

For this consider the general cubic equation as discussed below;

First denote;

$$L = k(\alpha_1, \alpha_2, \alpha_3)$$

$$F = k(a_1, a_2, a_3)$$

$$K = F(\alpha_i)$$

The polynomial will be denoted as;

$$\begin{aligned}f(x) &= x^3 - a_1x^2 + a_2x - a_3 \\ &= (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)\end{aligned}$$

This is irreducible in  $F[x]$

The symmetric group  $S_3$  acts on  $L$  by permuting  $\alpha_i$  and fixes the field  $F$ .

The field  $K = F(\alpha_1)$  is generated over  $F$  by the element  $\alpha_1$ , which is a root of the irreducible polynomial  $f(x) \in F[x]$

So;

$$K \cong F[t]/(f(t))$$

This has basis  $\{1, \alpha, \alpha^2\}$  over  $F$

Where;

$$\alpha_1 = 1$$

$$\alpha_2 = \alpha$$

$$\alpha_3 = \alpha^2$$

That is;

$$\dim_F K = 3$$

Now consider  $f(x)$  is a polynomial in  $K[x]$

It has a linear factor corresponding to the root  $\alpha_1 \in K$ , so, it factors as;

$$f(x) = (x - \alpha_1)g(x)$$

Where,  $g(x)$  is a quadratic polynomial in  $K[x]$

That is;

$$g(x) = (x - \alpha_2)(x - \alpha_3)$$

Substituting the values;

$$\begin{aligned} g(x) &= (x - \alpha)(x - \alpha^2) \\ &= x(x - \alpha^2) - \alpha(x - \alpha^2) \\ &= x^2 - x\alpha^2 - \alpha x + \alpha^3 \\ &= x^2 - x(\alpha^2 - \alpha) + \alpha^3 \end{aligned}$$

In particular;

$$g(x) = x^2 - x(\alpha_1^2 - \alpha_1) + \alpha_1^3$$

Therefore, the quadratic polynomial  $q(x)$  that appears in  $f(x) = (x - \alpha_1)g(x)$  explicitly in terms of  $\alpha_1$  and the coefficients of  $f$  is  $\boxed{g(x) = x^2 - x(\alpha_1^2 - \alpha_1) + \alpha_1^3}$

4. a



Consider the provided statement to check  $g(x)$  have a root in  $K$ .

As it is given that  $g(x) = x^3 + x + 1$  and  $\alpha$  is a root of the polynomial such that  $K = \mathbb{Q}(\alpha)$ .

Let  $f(x) = x^3 + 2x + 1$  and from the rational root test, it is find that  $f(x), g(x)$  both are irreducible in  $\mathbb{Q}[x]$ . Then from 16.2.8, if  $h(x) = x^3 + px + q$  is a polynomial function then their discriminants is provided as  $D = -4p^3 - 27q^2$ .

[Comment](#)

Step 2 of 3 ^

Therefore, the discriminant of the polynomial functions  $g(x) = x^3 + x + 1$  where  $p = 1, q = 1$  is as below:

$$\begin{aligned} D &= -4(1)^3 - 27(1)^2 \\ &= -4 - 27 \\ &= -31 \end{aligned}$$

The discriminant of the polynomial functions  $f(x) = x^3 + 2x + 1$  where  $p = 2, q = 1$  is as below:

$$\begin{aligned} D &= -4(2)^3 - 27(1)^2 \\ &= -4(8) - 27 \\ &= -32 - 27 \\ &= -59 \end{aligned}$$

Since, both obtained discriminant are not squares.

From the theorem 16.8.5, as their splitting fields  $L_f, L_g$  have degree 6 with Galois group over  $\mathbb{Q}$ . If  $g(x)$  has a root in  $K$  then it splits completely in  $K$  by the splitting theorem. Since both have degree 6 over  $\mathbb{Q}$  therefore  $L_f = L_g$ . As the Galois group is  $S_3$  there should be one intermediate field of degree 2 over  $\mathbb{Q}$  but the discriminant of polynomial functions are  $\mathbb{Q}(\sqrt{-59})$  and  $\mathbb{Q}(\sqrt{-31})$  therefore there is **contradiction**.

5. a

A cubic polynomial is a polynomial of degree 3. A closed-form solution known as the cubic formula exists for the solution of an arbitrary cubic equation.

[Comment](#)

Step 2 of 3 ^

Let  $\alpha_i$  be the roots of the cubic polynomial;

$$f(x) = x^3 + px + q$$

To find: a formula for a second root  $\alpha_2$  in terms of the element  $\alpha_1, \delta$  and the coefficients of  $f$

Let the root of  $x^3 + px + q$  be  $\alpha_1, \alpha_2, \alpha_3$

Then,

$$\alpha_1 + \alpha_2 + \alpha_3 = 0$$

Further;

$$\alpha_3 = -(\alpha_1 + \alpha_2)$$

And;

$$\alpha_i^3 = (-p\alpha_i + q)$$

Also;

$$\begin{aligned} q &= \alpha_1\alpha_2\alpha_3 \\ &= -(\alpha_1^2\alpha_2 + \alpha_1\alpha_2^2) \end{aligned}$$

Solve the above expression further;

$$\begin{aligned}\delta &= (\alpha_1 - \alpha_2)(2\alpha_1 + \alpha_2)(\alpha_1 + 2\alpha_2) \\ &= 2(\alpha_1^3 - \alpha_2^3) + 3(\alpha_1^2\alpha_2 - \alpha_1\alpha_2^2) \\ &= 2p(\alpha_1 - \alpha_2) + (\alpha_1^2\alpha_2 - \alpha_1\alpha_2^2)\end{aligned}$$

Here, both  $q$  and  $\delta$  are quadratic in  $\alpha_2$ , further eliminate the quadratic terms;

$$\delta - 3q = 2p(\alpha_1 - \alpha_2) + 6\alpha_1^2\alpha_2$$

Then solve for;

$$\begin{aligned}\delta - 3q &= 2p\alpha_1 - 2p\alpha_2 + 6\alpha_1^2\alpha_2 \\ \delta - 3q - 2p\alpha_1 &= (-2p + 6\alpha_1^2)\alpha_2 \\ \alpha_2 &= \frac{\delta - 3q - 2p\alpha_1}{6\alpha_1^2 - 2p}\end{aligned}$$

Therefore, a formula for a second root  $\alpha_2$  in terms of the element  $\alpha_1, \delta$  and the

coefficients of  $f$  is  $\alpha_2 = \frac{\delta - 3q - 2p\alpha_1}{6\alpha_1^2 - 2p}$

## Section 9

1. a

Symmetric group is defined as  $S_n$  on a finite set of groups whose elements are all the permutation operations.

[Comment](#)

Step 2 of 4 ^

Let  $K$  be a Galois extension of  $F$  whose Galois group is the symmetric group  $S_4$

To show: which integers occur as degree of elements of  $K$  over  $F$

For the proof first consider the permutations as,  $\sigma$

That is, associate each  $\sigma$  in the Galois group of  $f(T)$  over  $K$  its permutation on the roots of  $f(T)$

Consider these permutations as;

$$r_1, r_2, r_3, r_4$$

Now, these  $r_1, r_2, r_3, r_4$  will be homomorphism from the Galois group to  $S_4$

This homomorphism is injective since its kernel is trivial

An element of the Galois group that fixes each  $r_i$  is the identity on the splitting field.

Further two different choices for indexing the roots of  $f(T)$  can lead to different subgroups of  $S_4$  but they will be conjugate subgroups.

Like;

$$\begin{pmatrix} 1234 \\ 2431 \end{pmatrix} = (124)$$

This is the permutation turning one indexing of the roots into the other.

Now, the Galois group of  $f(T)$  over  $K$  does not have a canonical embedding into  $S_4$

That is image in  $S_4$  is well-defined up to an overall conjugation.

[Comment](#)

Step 4 of 4 ^

For a root  $r$  of  $f(T)$  in  $K$ ;

$$[K(r):K] = n$$

This is a factor of the degree of the splitting field over  $K$

**Therefore, the value of the integer which occur as degree of elements of  $K$  over  $F$  will be the degree of the splitting field over  $K$**

2. a

Nested radical expression is defined as the radical expression which contains a square root sign and then contains another radical expression with a square root.

[Comment](#)

Step 2 of 4 ^

To write: the element  $\alpha + \alpha'$  as a nested square root and what other nested square roots does  $K$  contains

For this consider  $\alpha$  to be a nested square root such that;

$$\alpha = \sqrt{a + \sqrt{b}}$$

Where  $a, b$  are constants

Now,  $\alpha'$  is defined as the conjugate thus;

$$\alpha' = \sqrt{a - \sqrt{b}}$$

The nested form can be written as;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

That is, consider;

$$A = \alpha + \alpha'$$

Here, consider;

$$a = 4$$

$$b = 5$$

Now, squaring both sides;

$$\begin{aligned} (A)^2 &= (\alpha + \alpha')^2 \\ &= \left( \sqrt{a+\sqrt{b}} + \sqrt{a-\sqrt{b}} \right)^2 \\ &= \left( \sqrt{4+\sqrt{5}} \right)^2 + 2 \left( \sqrt{4+\sqrt{5}} \right) \left( \sqrt{4-\sqrt{5}} \right) + \left( \sqrt{4-\sqrt{5}} \right)^2 \\ &= 4 + \sqrt{5} + 2 \sqrt{(4+\sqrt{5})(4-\sqrt{5})} + 4 - \sqrt{5} \\ &= 8 + 2\sqrt{16-5} \\ &= 8 + 2\sqrt{11} \end{aligned}$$

Thus,

$$A = \sqrt{8+2\sqrt{11}}$$

Therefore, the value for  $\alpha + \alpha'$  as nested square root will be  $\boxed{\sqrt{8+2\sqrt{11}}}$

And, consider;

$$B = \alpha - \alpha'$$

Now, squaring both sides;

$$\begin{aligned} (B)^2 &= (\alpha - \alpha')^2 \\ &= \left( \sqrt{a+\sqrt{b}} - \sqrt{a-\sqrt{b}} \right)^2 \\ &= \left( \sqrt{4+\sqrt{5}} \right)^2 - 2 \left( \sqrt{4+\sqrt{5}} \right) \left( \sqrt{4-\sqrt{5}} \right) + \left( \sqrt{4-\sqrt{5}} \right)^2 \\ &= 4 + \sqrt{5} - 2 \sqrt{(4+\sqrt{5})(4-\sqrt{5})} + 4 - \sqrt{5} \\ &= 8 - 2\sqrt{16-5} \\ &= 8 - 2\sqrt{11} \end{aligned}$$

Thus,

$$B = \sqrt{8-2\sqrt{11}}$$

Now, let  $K$  be a splitting field of  $f$  such that;

$$F \subset F(\alpha) \subset F(\alpha, \alpha')$$

And;

$$F(\alpha, \alpha') = K$$

Therefore, over the splitting field  $K$  the nested square root will be of the form  $\pm\alpha, \pm\alpha'$  where  $\alpha$  is of the form  $\sqrt{a+\sqrt{b}}$

Therefore, the nested square roots in  $K$  will be of the form  $\boxed{\pm\alpha, \pm\alpha', \pm A, \pm B}$

3. a

Rational numbers are those numbers which are of the form of  $\frac{p}{q}$  and  $q \neq 0$  where, both  $p, q \in \mathbb{Z}$ .

[Comment](#)

Step 2 of 3 ^

To show: whether  $\sqrt{4+\sqrt{7}}$  be written in the form  $\sqrt{a} + \sqrt{b}$ , with rational numbers  $a$  and  $b$

For this consider;

$$\sqrt{a} + \sqrt{b} = \sqrt{4+\sqrt{7}}$$

Now, squaring both sides;

$$\begin{aligned} (\sqrt{a} + \sqrt{b})^2 &= (\sqrt{4+\sqrt{7}})^2 \\ (\sqrt{a})^2 + 2(\sqrt{a})(\sqrt{b}) + (\sqrt{b})^2 &= 4 + \sqrt{7} \\ a + 2\sqrt{ab} + b &= 4 + \sqrt{7} \\ (a+b) + 2\sqrt{ab} &= 4 + \sqrt{7} \end{aligned}$$

Now, comparing the terms;

$$\begin{aligned} a + b &= 4 \\ 2\sqrt{ab} &= \sqrt{7} \end{aligned}$$

Now, squaring both sides;

$$\begin{aligned} (2\sqrt{ab})^2 &= (\sqrt{7})^2 \\ 4ab &= 7 \end{aligned}$$

Now, first take the equation;

$$\begin{aligned} a + b &= 4 \\ a &= 4 - b \end{aligned}$$

Substitute this value in  $4ab = 7$ ;

$$\begin{aligned} 4(4-b)b &= 7 \\ (16-4b)b &= 7 \\ 16b-4b^2 &= 7 \\ 4b^2-16b+7 &= 0 \end{aligned}$$

Now, finding the values by the method of middle term splitting;

$$\begin{aligned} 4b^2 - 2b - 14b + 7 &= 0 \\ 2b(2b-1) - 7(2b-1) &= 0 \\ (2b-7)(2b-1) &= 0 \end{aligned}$$

That is;

$$\begin{aligned} 2b-7=0, 2b-1 &= 0 \\ b &= \frac{7}{2}, b = \frac{1}{2} \end{aligned}$$

$$\text{When, } b = \frac{7}{2}$$

$$\text{Then, } a = \frac{1}{2}$$

$$\text{And, when } b = -\frac{7}{2}$$

$$\text{Then } a = \frac{15}{2}$$

Thus the  $\sqrt{4+\sqrt{7}}$  can be written in the form  $\sqrt{a} + \sqrt{b}$ , where;

$$\boxed{a = \frac{1}{2}, b = \frac{7}{2} \text{ and } a = \frac{15}{2}, b = -\frac{7}{2}}$$

4. a

An irreducible polynomial is defined as a non-constant polynomial that cannot be factored into the product of two non-constant polynomials

[Comment](#)

Step 2 of 6 ^

a.

To prove: that the polynomial  $x^4 - 8x^2 + 11$  is irreducible over  $\mathbb{Q}$  in two ways

First way;

Consider the given polynomial:

$$x^4 - 8x^2 + 11$$

Now, there will be four factors of the given polynomial as the degree is 4, where, the roots are;

$$-\sqrt{4-\sqrt{5}}, \sqrt{4-\sqrt{5}}, -\sqrt{4+\sqrt{5}}, \sqrt{4+\sqrt{5}}$$

Since,  $\mathbb{Q}$  is the set of rational number.

So, clearly the roots obtained for the polynomial does not lie in the given field, that is  $\mathbb{Q}$

**Therefore, the polynomial  $x^4 - 8x^2 + 11$  is irreducible over  $\mathbb{Q}$**

Second way;

Consider the given polynomial:

$$x^4 - 8x^2 + 11$$

Since, from the above part clearly the given polynomial is not factorizable.

Thus, this means that the polynomial is a prime integral and also, the field  $\mathbb{Q}$  forms a totally ordered field under addition and multiplication.

Thus, the field of rational numbers is integral domain under addition and multiplication.

Let the polynomial to be a prime integral, that is denote it by  $p$ , such that;

$$p = 1 \times p$$

Where;

$$a = 1$$

$$b = p$$

And;

$$a < p$$

$$b < p$$

Now, by using the theorem which states that;

In an integral domain  $R$ , a prime element is irreducible.

Here,

$$\mathbb{Q} \subseteq R$$

So, clearly from the theorem, **the polynomial  $x^4 - 8x^2 + 11$  is irreducible over  $\mathbb{Q}$**

b.

To prove: that the polynomial  $x^4 - 3x^2 + 9$  is irreducible over  $\mathbb{Q}$  in two ways

First way;

Consider the given polynomial:

$$x^4 - 3x^2 + 9$$

Now, there will be four factors of the given polynomial as the degree is 4, where, the roots are;

$$-\frac{1}{2}i(\sqrt{3}-3i), \frac{1}{2}(3-i\sqrt{3}), \frac{1}{2}i(\sqrt{3}+3i), \frac{1}{2}(3+i\sqrt{3})$$

Since,  $\mathbb{Q}$  is the set of rational number.

So, clearly the roots obtained for the polynomial does not lie in the given field, that is  $\mathbb{Q}$

**Therefore, the polynomial  $x^4 - 3x^2 + 9$  is irreducible over  $\mathbb{Q}$**



Second way;

Consider the given polynomial:

$$x^4 - 3x^2 + 9$$

Since, from the above part clearly the given polynomial is not factorizable.

Thus, this means that the polynomial is a prime integral and also, the field  $\mathbb{Q}$  forms a totally ordered field under addition and multiplication.

Thus, the field of rational numbers is integral domain under addition and multiplication.

Let the polynomial to be a prime integral, that is denote it by  $p$ , such that;

$$p = 1 \times p$$

Where;

$$a = 1$$

$$b = p$$

And;

$$a < p$$

$$b < p$$

Now, by using the theorem which states that;

In an integral domain  $R$ , a prime element is irreducible.

Here,

$$\mathbb{Q} \subset R$$

So, clearly from the theorem, **the polynomial  $x^4 - 3x^2 + 9$  is irreducible over  $\mathbb{Q}$**

c.

To determine: all intermediate fields when  $K$  is the splitting field of  $x^4 - 8x^2 + 11$

Since, the given polynomial is of the form;

$$ax^4 + cx^2 + e = 0$$

Then, the determinant is of the form of;

$$\Delta = 256a^3e^3 - 128a^2c^2e^2 + 16ac^4e$$

Comparing the given polynomial;

$$a = 1$$

$$c = -8$$

$$e = 11$$

Now, substitute to find the determinant;

$$\begin{aligned} \Delta &= 256a^3e^3 - 128a^2c^2e^2 + 16ac^4e \\ &= 256(1)^3(11)^3 - 128(1)^2(-8)^2(11)^2 + 16(1)(-8)^4(11) \\ &= 70400 \end{aligned}$$

From the first part the polynomial is irreducible and the determinant is positive but not a square.

Thus, the intermediate polynomial will become  $S_4$

**Therefore, the intermediate fields when  $K$  is the splitting field of  $x^4 - 8x^2 + 11$  is  $S_4$**

## 5. a

Nested radical expression is defined as the radical expression which contains a square root sign and then contains another radical expression with a square root.

Consider a nested square root  $\alpha = \sqrt{r + \sqrt{t}}$  with  $r, t$  in a field  $F$ . Assume that  $\alpha$  has degree 4 over  $F$ , let  $f$  be the irreducible polynomial of  $\alpha$  over  $F$ , and let  $K$  be a splitting field of  $f$  over  $F$ .

a.

To compute: the irreducible polynomial for  $\alpha$  over  $F$  and prove that  $G(K/F)$  is one of the groups  $D_4, C_4$  or  $D_2$ .

First consider  $\alpha_1^2 \in F$  and  $\alpha_2^2 \in F(\alpha_1)$

Here,  $\alpha_2^2$  is not an element of  $F$  because if it were,

$$K^{[1, (12)(34)]} = F(\alpha_1)(\alpha_2)$$

This would be the splitting field of  $(x^2 - \alpha_1^2)(x^2 - \alpha_2^2)$ , but this would violate the non-normality of;

$$\{I, (12)(34)\} \subset D_4$$

And, since;

$$\begin{aligned} [F(\alpha_1) : F] &= 2 \\ \alpha_2^2 &\in F(\alpha_1) - F \end{aligned}$$

Thus, conclude that;

$$F(\alpha_1) = F[\alpha_2^2]$$

Thus,

$$F(\alpha_1, \alpha_2) = F[\alpha_2]$$

Therefore, the irreducible polynomial for  $\alpha$  over  $F$  is of the form  $F[\sqrt{\alpha}]$

Now, since;

$$[F[\alpha_2^2] : F] = 2$$

Then,  $\alpha_2^2$  satisfies an irreducible polynomial  $x^2 + bx + c \in F[x]$  and so,  $\alpha_2$  satisfies the polynomial;

$$x^4 + bx^2 + c \in F[x]$$

This must be irreducible since;

$$[F[\alpha_2] : F] = 4$$

Here,  $K$  is the splitting field of;

$$x^4 + bx^2 + c$$

Since, the polynomial cannot split in;

$$F[\alpha_2] = K^{[1, (12)(34)]}$$

Since,  $\{I, (12)(34)\}$  is not normal in  $D_4$

Therefore,  $G(K/F) = D_4$

b.

To explain: how to determine the Galois group in terms of the element  $r^2 - t$

Since, the element is of the form  $r^2 - t$  then the use the result which states that;

Let  $f(x)$  be an irreducible quartic in  $\mathbb{Q}[x]$ . If  $\text{disc } f > 0$  then  $G_f = \mathbb{Z}/4\mathbb{Z}$  otherwise  $G_f = D_4$

Clearly for the element of the form  $r^2 - t$

The determinant that is the;

$$\text{disc } f > 0$$

Therefore, the Galois group in terms of the element  $r^2 - t$  is of the form of  $G = \mathbb{Z}/4\mathbb{Z}$

c.

Assume that Galois group of  $K/F$  is the dihedral group  $D_4$

To determine: generators for all intermediate fields  $F \subset L \subset K$

The Dihedral group  $D_4$  can be constructed as follows;

$$\begin{aligned} D_4 &\supset \{I, (12)(34), (13)(24), (14)(23)\} \\ &\supset \{I, (12)(34)\} \\ &\supset \{I\} \end{aligned}$$

Where each subgroup is normal in  $D_4$

Now, by the construction theorem the, another corresponding ladder of fields can be represented as given below:

$$\begin{aligned} K^{D_4} &= F \\ &\subset K^{\{I, (12)(34), (13)(24), (14)(23)\}} \\ &= F(\alpha_1)(\alpha_2) \\ &\subset K \end{aligned}$$

Therefore, the generators for all intermediate fields  $F \subset L \subset K$  of Dihedral group  $D_4$  will

be  $\boxed{\{I, (12)(34), (13)(24), (14)(23)\}}$

6. a

Consider the provided statement to compute the discriminant of the given quartic polynomial and also determine its Galois group over  $\mathbb{Q}$ .

As provided quartic polynomial is  $x^4 + 1$ ,

The discriminant  $D_n$  for the polynomial function  $x^n + px + q$  whose degree is greater than or equal to 2 is given as below:

$$D_n = (-1)^{\frac{n(n-1)}{2}} (n-1)^{n-1} p^n + (-1)^{\frac{n(n-1)}{2}} n^n q^{n-1}$$

Therefore discriminant of  $x^4 + 1$  is provided as below where  $n = 4, p = 0, q = 1$ . Then from the theorem 16.2.4, the discriminant is  $D_4 = -27p^4 + 256q^3$ . Therefore,

$$\begin{aligned} D_4 &= -27p^4 + 256q^3 \\ &= -27(0)^4 + 256(1)^3 \\ &= 256 \end{aligned}$$

From the proposition 19.9.5, it implies that the Galois group  $G$  is  $A_4$  or  $D_2$ . The roots of the polynomial function  $x^4 + 1$  are as below:

$$\begin{aligned} \alpha_1 &= \frac{1}{\sqrt{2}}(1+i) \\ \alpha_2 &= -\frac{1}{\sqrt{2}}(1+i) \\ \alpha_3 &= \frac{1}{\sqrt{2}}(1-i) \\ \alpha_4 &= -\frac{1}{\sqrt{2}}(1-i) \end{aligned}$$

The values of  $\beta_1, \beta_2, \beta_3$  are as follows  $\beta_1 = 0, \beta_2 = 2, \beta_3 = -2$  and  $g(x) = x(x-2)(x+2)$ . Therefore, from proposition 19.6.8 the value of  $G = D_2$ .

7. a

Nested radical expression is defined as the radical expression which contains a square root sign and then contains another radical expression with a square root.

Consider that an extension field  $K/F$  has the form  $K = F(\sqrt{a}, \sqrt{b})$

To determine: all nested square roots  $\sqrt{r + \sqrt{t}}$  that are in  $K$ , with  $r$  and  $t$  in  $F$

Here the mapping is defined as;

$$\begin{aligned} K &= F(\sqrt{a}, \sqrt{b}) \\ &= \sqrt{r + \sqrt{t}} \end{aligned}$$

Now, this can be denested as follows;

$$\begin{aligned} \sqrt{r + \sqrt{t}} &= \sqrt{(\sqrt{r + \sqrt{t}})^2} \\ &= \sqrt{(c + d) + 2\sqrt{cd}} \\ &= \sqrt{a + \sqrt{b}} \end{aligned}$$

Where;

$$\begin{aligned} r + t &= a \\ rt &= \frac{b}{4} \end{aligned}$$

This means that  $r$  and  $t$  are therefore, the solutions of  $x^2 - ax + \frac{b}{4} = 0$

Similarly, it holds for;

$$\sqrt{a - \sqrt{b}} = \sqrt{r} - \sqrt{t}$$

If the roots are rational for  $\sqrt{w + \sqrt{x} + \sqrt{y} + \sqrt{z}}$ ;

$$\begin{aligned} \sqrt{a + \sqrt{b} + \sqrt{c}} &= \sqrt{(\sqrt{a + \sqrt{b} + \sqrt{c}})^2} \\ &= \sqrt{(a + b + c) + \sqrt{4ab} + \sqrt{4ac} + \sqrt{4bc}} \\ &= \sqrt{a + \sqrt{b}} \end{aligned}$$

Where,

$$\begin{aligned} a + b + c &= w \\ ab + ac + bc &= \frac{x + y + z}{4} \\ abc &= \frac{\sqrt{xyz}}{8} \end{aligned}$$

Thus,  $a, b, c$  are solutions of the cubic equation;

$$u^3 - wu^2 + \frac{x + y + z}{4}u - \frac{\sqrt{xyz}}{8} = 0$$

**Therefore, the nested form depends upon the degree of the function.**

## Method 2.

**Given:**  $K$  is a field extension of  $F$  given by the form  $K = F(\sqrt{a}, \sqrt{b})$ .

**Solution:** We will determine all nested square roots  $\sqrt{r + \sqrt{t}}$  that are in  $K$  and  $r, t$  are in  $F$ .  
Let us consider the following cases.

**Case-1:**  $t$  is a perfect square in  $F$ . In this case it is all trivial.

**Case-2:**  $t$  is not a perfect square in  $F$ .

Now notice that  $\sqrt{t} \in F(\sqrt{r + \sqrt{t}})$ . And therefore we have

$$\sqrt{t} \in F(\sqrt{a}, \sqrt{b}).$$

This follows that,  $t$  is of the following form

$$t = k^2a \text{ or } t = k^2b \text{ or } t = k^2ab, \text{ where } k \in F.$$

Now notice that  $k^2a$  and  $k^2b$  are symmetric one, so we can consider the forms as

$$t = k^2a \text{ or } t = k^2ab.$$

So we have

$$F(\sqrt{a}, \sqrt{b}) = \sqrt{a}, \sqrt{ab}.$$

Let us now look at the equation for  $m, n, p, q \in F$

$$(m + n\sqrt{a} + p\sqrt{b} + q\sqrt{ab})^2 = r + k\sqrt{a}.$$

Equating the coefficients we have

$$mq + np = 0 \tag{1}$$

$$mp + anq = 0. \tag{2}$$

Let us consider that all four variables in the above equations are non-zero.

Now from (1) we have  $q = -\frac{np}{m}$ .

Now substituting this into the second equation we have

$$m^2 = an^2.$$

And this is possible only if  $m = n = 0$ .

So basically we have either  $m = n = 0$  or  $p = q = 0$ .

Thus a square root of  $r + k\sqrt{a}$  must have one of the following form

$$m + n\sqrt{a} \text{ or } \sqrt{b}(m + n\sqrt{a}).$$

Now notice that

$$(m + n\sqrt{a})^2 = (m^2 + an^2) + 2mn\sqrt{a}$$

and

$$(\sqrt{b}(m + n\sqrt{a}))^2 = b(m^2 + an^2) + 2mnb\sqrt{a}.$$

This follows that  $\sqrt{r + \sqrt{t}} \in F(\sqrt{a}, \sqrt{b})$ , where  $r = m^2 + an^2$  and  $t = 4am^2n^2$ , for  $m, n \in F$ .  
Or we can assume  $r = b(m^2 + an^2)$  and  $t = 4ab^2m^2n^2$ .

By the similar argument we can show that the elements of the form  $m + n\sqrt{b}$  and  $\sqrt{a}(m + n\sqrt{b})$  can equal a nested radical of the form  $\sqrt{r + \sqrt{t}}$ .  
This completes the solution.

## Result

3 of 3

We have determine all nested square roots  $\sqrt{r + \sqrt{t}}$  that are in  $K$  and  $r, t$  are in  $F$ .

## 8. a

To determine: whether or not the following nested radicals can be written in the form of unnested square roots if yes then find its expression

a.

Consider the given expression;

$$\sqrt{2 + \sqrt{11}}$$

Let  $\alpha$  be the nested square root, that is;

$$\alpha = \sqrt{2 + \sqrt{11}}$$

To determine the irreducible polynomial for  $\alpha$  over  $F$ , then its roots will be  $\pm\alpha, \pm\alpha'$

Here;

$$\alpha' = \sqrt{2 - \sqrt{11}}$$

To find the unnested part find;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

For this, consider;

$$\begin{aligned} a &= \alpha + \alpha' \\ &= \sqrt{2 + \sqrt{11}} + \sqrt{2 - \sqrt{11}} \end{aligned}$$

Squaring both sides;

$$\begin{aligned} a^2 &= \left( \sqrt{2 + \sqrt{11}} + \sqrt{2 - \sqrt{11}} \right)^2 \\ &= \left( \sqrt{2 + \sqrt{11}} \right)^2 + \left( \sqrt{2 - \sqrt{11}} \right)^2 + 2 \left( \sqrt{2 + \sqrt{11}} \right) \left( \sqrt{2 - \sqrt{11}} \right) \\ &= 2 + \sqrt{11} + 2 - \sqrt{11} + 2 \left( \sqrt{(2 + \sqrt{11})(2 - \sqrt{11})} \right) \\ &= 4 + 2 \left( \sqrt{4 - 11} \right) \\ &= 4 + \sqrt{-8} \end{aligned}$$

This is not possible as there is a negative integer in the square root. Thus the unnested form is not possible.



**b.**

Consider the given expression;

$$\sqrt{10+5\sqrt{2}}$$

Let  $\alpha$  be the nested square root, that is;

$$\alpha = \sqrt{10+5\sqrt{2}}$$

To determine the irreducible polynomial for  $\alpha$  over  $F$ , then its roots will be  $\pm\alpha, \pm\alpha'$

Here;

$$\alpha' = \sqrt{10-5\sqrt{2}}$$

To find the unnested part find;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

For this, consider;

$$\begin{aligned} a &= \alpha + \alpha' \\ &= \sqrt{10+5\sqrt{2}} + \sqrt{10-5\sqrt{2}} \end{aligned}$$

Squaring both sides;

$$\begin{aligned} a^2 &= \left( \sqrt{10+5\sqrt{2}} + \sqrt{10-5\sqrt{2}} \right)^2 \\ &= \left( \sqrt{10+5\sqrt{2}} \right)^2 + \left( \sqrt{10-5\sqrt{2}} \right)^2 + 2\left( \sqrt{10+5\sqrt{2}} \right)\left( \sqrt{10-5\sqrt{2}} \right) \\ &= 10+5\sqrt{2} + 10-5\sqrt{2} + 2\left( \sqrt{(10+5\sqrt{2})(10-5\sqrt{2})} \right) \\ &= 20 + 2\left( \sqrt{100-50} \right) \\ &= 20 + 2\sqrt{50} \end{aligned}$$

And;

$$\begin{aligned} b &= \alpha - \alpha' \\ &= \sqrt{10+5\sqrt{2}} - \sqrt{10-5\sqrt{2}} \end{aligned}$$

Squaring both sides;

$$\begin{aligned} b^2 &= \left( \sqrt{10+5\sqrt{2}} - \sqrt{10-5\sqrt{2}} \right)^2 \\ &= \left( \sqrt{10+5\sqrt{2}} \right)^2 + \left( \sqrt{10-5\sqrt{2}} \right)^2 - 2\left( \sqrt{10+5\sqrt{2}} \right)\left( \sqrt{10-5\sqrt{2}} \right) \\ &= 10+5\sqrt{2} + 10-5\sqrt{2} - 2\left( \sqrt{(10+5\sqrt{2})(10-5\sqrt{2})} \right) \\ &= 20 - 2\left( \sqrt{100-50} \right) \\ &= 20 - 2\sqrt{50} \end{aligned}$$

**This one is also not possible as unnesting is not being processed.**

**c.**

Consider the given expression;

$$\sqrt{11+6\sqrt{2}}$$

Let  $\alpha$  be the nested square root, that is;

$$\alpha = \sqrt{11+6\sqrt{2}}$$

To determine the irreducible polynomial for  $\alpha$  over  $F$ , then its roots will be  $\pm\alpha, \pm\alpha'$

Here;

$$\alpha' = \sqrt{11-6\sqrt{2}}$$

To find the unnested part find;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

For this, consider;

$$\begin{aligned} a &= \alpha + \alpha' \\ &= \sqrt{11+6\sqrt{2}} + \sqrt{11-6\sqrt{2}} \end{aligned}$$

Squaring both sides;

$$\begin{aligned} a^2 &= \left( \sqrt{11+6\sqrt{2}} + \sqrt{11-6\sqrt{2}} \right)^2 \\ &= \left( \sqrt{11+6\sqrt{2}} \right)^2 + \left( \sqrt{11-6\sqrt{2}} \right)^2 + 2 \left( \sqrt{11+6\sqrt{2}} \right) \left( \sqrt{11-6\sqrt{2}} \right) \\ &= 11+6\sqrt{2} + 11-6\sqrt{2} + 2 \left( \sqrt{(11+6\sqrt{2})(11-6\sqrt{2})} \right) \\ &= 22 + 2 \left( \sqrt{121-72} \right) \\ &= 22 + 2\sqrt{49} \\ &= 22 + 2(7) \\ &= 22 + 14 \\ &= 36 \end{aligned}$$

That is;

$$\begin{aligned} a &= \sqrt{36} \\ &= 6 \end{aligned}$$

Next, consider;

$$\begin{aligned} b &= \alpha - \alpha' \\ &= \sqrt{11+6\sqrt{2}} - \sqrt{11-6\sqrt{2}} \end{aligned}$$

Squaring both sides;

$$\begin{aligned} b^2 &= \left( \sqrt{11+6\sqrt{2}} - \sqrt{11-6\sqrt{2}} \right)^2 \\ &= \left( \sqrt{11+6\sqrt{2}} \right)^2 + \left( \sqrt{11-6\sqrt{2}} \right)^2 - 2 \left( \sqrt{11+6\sqrt{2}} \right) \left( \sqrt{11-6\sqrt{2}} \right) \\ &= 11+6\sqrt{2} + 11-6\sqrt{2} - 2 \left( \sqrt{(11+6\sqrt{2})(11-6\sqrt{2})} \right) \\ &= 22 - 2 \left( \sqrt{121-72} \right) \\ &= 22 - 2\sqrt{49} \\ &= 22 - 2(7) \\ &= 22 - 14 \\ &= 8 \end{aligned}$$

That is;

$$b = \sqrt{8}$$

So, the unnested form will be;

$$\begin{aligned} \alpha &= \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2} \\ &= \frac{6 + \sqrt{8}}{2} \\ &= \frac{6 + 2\sqrt{2}}{2} \\ &= 3 + \sqrt{2} \end{aligned}$$

Therefore, the unnested form will be  $\boxed{3 + \sqrt{2}}$

Now, finding an expression for this unnested square root term;

$$f(x) = (x - \alpha)(x + \alpha)(x - \alpha')(x + \alpha')$$

Substituting the values;

$$\begin{aligned} f(x) &= \left(x - \sqrt{11 + 6\sqrt{2}}\right)\left(x + \sqrt{11 + 6\sqrt{2}}\right)\left(x - \sqrt{11 - 6\sqrt{2}}\right)\left(x + \sqrt{11 - 6\sqrt{2}}\right) \\ &= \left(x^2 - \left(\sqrt{11 + 6\sqrt{2}}\right)^2\right)\left(x^2 - \left(\sqrt{11 - 6\sqrt{2}}\right)^2\right) \\ &= \left(x^2 - (11 + 6\sqrt{2})\right)\left(x^2 - (11 - 6\sqrt{2})\right) \\ &= (x^2)^2 - x^2(11 - 6\sqrt{2}) - x^2(11 + 6\sqrt{2}) + (11 + 6\sqrt{2})(11 - 6\sqrt{2}) \\ &= x^4 - 11x^2 + 6\sqrt{2}x^2 - 11x^2 - 6\sqrt{2}x^2 + 121 - 72 \\ &= x^4 - 22x^2 + 49 \end{aligned}$$

Therefore the required expression is  $\boxed{x^4 - 22x^2 + 49}$

d.

Consider the given expression;

$$\sqrt{6 + \sqrt{11}}$$

Let  $\alpha$  be the nested square root, that is;

$$\alpha = \sqrt{6 + \sqrt{11}}$$

To determine the irreducible polynomial for  $\alpha$  over  $F$ , then its roots will be  $\pm\alpha, \pm\alpha'$

Here;

$$\alpha' = \sqrt{6 - \sqrt{11}}$$

To find the unnested part find;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

For this, consider;

$$\begin{aligned} a &= \alpha + \alpha' \\ &= \sqrt{6 + \sqrt{11}} + \sqrt{6 - \sqrt{11}} \end{aligned}$$

Squaring both sides;

$$\begin{aligned} a^2 &= \left(\sqrt{6 + \sqrt{11}} + \sqrt{6 - \sqrt{11}}\right)^2 \\ &= \left(\sqrt{6 + \sqrt{11}}\right)^2 + \left(\sqrt{6 - \sqrt{11}}\right)^2 + 2\left(\sqrt{6 + \sqrt{11}}\right)\left(\sqrt{6 - \sqrt{11}}\right) \\ &= 6 + \sqrt{11} + 6 - \sqrt{11} + 2\left(\sqrt{(6 + \sqrt{11})(6 - \sqrt{11})}\right) \\ &= 12 + 2\left(\sqrt{36 - 11}\right) \\ &= 12 + 2\left(\sqrt{25}\right) \\ &= 12 + 2(5) \quad \text{That is;} \\ &= 12 + 10 \\ &= 22 \\ a &= \sqrt{22} \end{aligned}$$

Next, consider;

$$b = \alpha - \alpha'$$

$$= \sqrt{6 + \sqrt{11}} - \sqrt{6 - \sqrt{11}}$$

Squaring both sides;

$$b^2 = \left( \sqrt{6 + \sqrt{11}} - \sqrt{6 - \sqrt{11}} \right)^2$$

$$= \left( \sqrt{6 + \sqrt{11}} \right)^2 + \left( \sqrt{6 - \sqrt{11}} \right)^2 - 2 \left( \sqrt{6 + \sqrt{11}} \right) \left( \sqrt{6 - \sqrt{11}} \right)$$

$$= 6 + \sqrt{11} + 6 - \sqrt{11} - 2 \left( \sqrt{(6 + \sqrt{11})(6 - \sqrt{11})} \right)$$

$$= 12 - 2 \left( \sqrt{36 - 11} \right)$$

$$= 12 - 2 \left( \sqrt{25} \right)$$

$$= 12 - 2(5) \quad \text{That is;}$$

$$= 12 - 10$$

$$= 22$$

$$b = \sqrt{22}$$

So, the unnested form will be;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

$$= \frac{\sqrt{22} + \sqrt{2}}{2}$$

Therefore, the unnested form will be  $\boxed{\frac{\sqrt{22} + \sqrt{2}}{2}}$

Now, finding an expression for this unnested square root term;

$$f(x) = (x - \alpha)(x + \alpha)(x - \alpha')(x + \alpha')$$

Substituting the values;

$$f(x) = (x - \sqrt{6 + \sqrt{11}})(x + \sqrt{6 + \sqrt{11}})(x - \sqrt{6 - \sqrt{11}})(x + \sqrt{6 - \sqrt{11}})$$

$$= \left( x^2 - \left( \sqrt{6 + \sqrt{11}} \right)^2 \right) \left( x^2 - \left( \sqrt{6 - \sqrt{11}} \right)^2 \right)$$

$$= \left( x^2 - (6 + \sqrt{11}) \right) \left( x^2 - (6 - \sqrt{11}) \right)$$

$$= (x^2)^2 - x^2(6 - \sqrt{11}) - x^2(6 + \sqrt{11}) + (6 + \sqrt{11})(6 - \sqrt{11})$$

$$= x^4 - 6x^2 + \sqrt{11}x^2 - 6x^2 - \sqrt{11}x^2 + 36 - 11$$

$$= x^4 - 12x^2 + 25$$

Therefore the required expression is  $\boxed{x^4 - 12x^2 + 25}$

e .

Consider the given expression;

$$\sqrt{11 + \sqrt{6}}$$

Let  $\alpha$  be the nested square root, that is;

$$\alpha = \sqrt{11 + \sqrt{6}}$$

To determine the irreducible polynomial for  $\alpha$  over  $F$ , then its roots will be  $\pm\alpha, \pm\alpha'$

Here;

$$\alpha' = \sqrt{11 - \sqrt{6}}$$

To find the unnested part find;

$$\alpha = \frac{(\alpha + \alpha') + (\alpha - \alpha')}{2}$$

For this, consider;

$$a = \alpha + \alpha'$$

$$= \sqrt{11 + \sqrt{6}} + \sqrt{11 - \sqrt{6}}$$

Squaring both sides;

$$\begin{aligned}
 a^2 &= \left( \sqrt{11+\sqrt{6}} + \sqrt{11-\sqrt{6}} \right)^2 \\
 &= \left( \sqrt{11+\sqrt{6}} \right)^2 + \left( \sqrt{11-\sqrt{6}} \right)^2 + 2 \left( \sqrt{11+\sqrt{6}} \right) \left( \sqrt{11-\sqrt{6}} \right) \\
 &= 11+\sqrt{6}+11-\sqrt{6}+2 \left( \sqrt{(11+\sqrt{6})(11-\sqrt{6})} \right) \\
 &= 11+2 \left( \sqrt{121-6} \right) \\
 &= 11+2\sqrt{115}
 \end{aligned}$$

Since, the square root has not been absorbed. Therefore, this cannot be written in the form unnested square roots.

9. a

The discriminant is defined as a symmetric function in the roots; it can be expressed in terms of the coefficients of the polynomial.

a.

To determine: the discriminant and the resolvent cubic of a polynomial of the form ;

$$f(x) = x^4 + rx + s$$

In general the determinant of the fourth order equation, that is;

$$a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0$$

Its discriminant is given by;

$$D_4 = \left[ \begin{aligned} &(a_1^2a_2^2a_3^2 - 4a_1^3a_3^3 - 4a_1^2a_2^3a_4 + 18a_1^3a_2a_3a_4 - 27a_1^4a_4^2 + 256a_0^3a_4^3) + \\ &a_0(-4a_2^3a_3^2 + 18a_1a_2a_3^3 + 16a_2^4a_4 - 80a_1a_2^2a_3a_4 - 6a_1^2a_3^2a_4 + 144a_1^2a_2a_4^2) + \\ &a_0^2(-27a_3^4 + 144a_2a_3^2a_4 - 128a_2^2a_4^2 - 192a_1a_3a_4^2) \end{aligned} \right]$$

Now, from the given expression  $f(x) = x^4 + rx + s$  comparing the coefficients;

$$a_4 = 1$$

$$a_3 = 0$$

$$a_2 = 0$$

$$a_1 = r$$

$$a_0 = s$$

Now, substituting these values in the formula for the discriminant;

$$\begin{aligned}
 D_4 &= \left[ (0-0-0+0-27r^4+256s^3) + s(0+0+0-0-0+0) + s^2(0+0-0-0) \right] \\
 &= -27r^4 + 256s^3
 \end{aligned}$$

Hence, the required discriminant is  $\boxed{-27r^4 + 256s^3}$

Further to find the cubic resolvent;

When  $f(x)$ , that is the function is a quartic defined as;

$$f(x) = x^4 + cx + d$$

With roots  $r_1, r_2, r_3, r_4$  then its cubic resolvent will be defined as;

$$R_3(x) = x^3 - 4dx - c^2$$

Now, comparing the values from the given equation,  $f(x) = x^4 + rx + s$ , that is;

$$c = r$$

$$d = s$$

Substituting this value in the cubic resolvent formula;

$$\begin{aligned}
 R_3(x) &= x^3 - 4(s)x - r^2 \\
 &= x^3 - 4sx - r^2
 \end{aligned}$$

Hence, the required cubic resolvent is  $\boxed{x^3 - 4sx - r^2}$

b.

To determine: the Galois groups of  $x^4 + 8x + 12$  and  $x^4 + 8x - 12$  over  $\mathbb{Q}$

First finding the Galois group of  $x^4 + 8x + 12$  over  $\mathbb{Q}$

From the above resolvent formula the resolvent cubic will be;

$$x^3 - 4(12)x + (8)^2 = x^3 - 48x + 64$$

This expression does not have rational roots and its discriminant will be;

$$\begin{aligned} -27 \times 8^4 + 256 \times (12)^3 &= 27(2^{14} - 2^{12}) \\ &= 81 \times 2^{12} \end{aligned}$$

Since, this is a perfect square

**Therefore, the Galois group is  $A_4$**

Now, finding the Galois group of  $x^4 + 8x - 12$  over  $\mathbb{Q}$

From the above resolvent formula the resolvent cubic will be;

$$x^3 - 4(-12)x + (8)^2 = x^3 + 48x + 64$$

And its discriminant will be;

$$\begin{aligned} -27 \times 8^4 + 256 \times (-12)^3 &= -27(2^{14} - 2^{12}) \\ &= -81 \times 2^{12} \\ &< 0 \end{aligned}$$

This implies the roots are not rational

**Therefore, the Galois group is  $S_4$**

c.

To show: that the roots of the polynomial  $x^4 + x - 5$  be constructed by ruler and compass

Roots are defined as the zeroes of the function at which the value of function gets vanish.

Now, consider the statement given below;

Suppose there are points;

$$\begin{aligned} P_1 &= (a_1, b_1), \dots, P_n \\ &= (a_n, b_n) \end{aligned}$$

Consider this in the real Cartesian plane.

Then it is possible to construct a point  $Q = (\alpha, \beta)$  with ruler and compass if and only if  $\alpha$  and  $\beta$  can be obtained from  $a_1, \dots, a_n, b_1, \dots, b_n$  by field operations and the solution of a finite number of successive linear and quadratic equations involving the square roots of positive real numbers.

This quartic equation is not possible to solve to obtain its roots. Therefore, from the above theorem it can be said that **it is not possible to construct it using the ruler and compass.**

10. a



Galois group is defined as the group of field automorphisms which true under any composition.

[Comment](#)

Step 2 of 3 ^

a.

To find: the possible Galois group of an irreducible quartic polynomial over  $\mathbb{Q}$  that has exactly two real roots.

For this, if  $f$  is irreducible then its Galois group acts transitively.

Now, the transitive subgroups of  $S_4$  are;

$S_4, A_4, D_4, C_4,$  and  $C_2 \times C_2$

If  $f$  has exactly two real roots, then it has two complex roots and complex conjugation is an element of its Galois group

That is the all the sets are of order 2

**Therefore, the possible Galois group of an irreducible quartic polynomial over  $\mathbb{Q}$  that has exactly two real roots is  $S_4, A_4, D_4, C_4,$  and  $C_2 \times C_2$**

b.

To find: the possible Galois groups over  $\mathbb{Q}$  of an irreducible quartic polynomial  $f(x)$  whose discriminant is negative.

For this consider  $F$  be a field of characteristic not equal to 2.

Denote the negative determinant by  $\Delta$

First look for the reducibility of  $p(x)$  over  $F(\sqrt{\Delta})$

If  $\Delta$  is not a square and the resolvent cubic is reducible, then;

$$\text{Gal}(p/f) = D_4$$

If and only if  $p(x)$  is irreducible in  $F(\sqrt{\Delta})$

Now, if  $\alpha$  is a root of  $p$  and has degree 8 over  $F$

But this implies that;

$$[F(\sqrt{\Delta}, \alpha) : F(\sqrt{\Delta})] = 4$$

From this it can be derived that  $p$  must be irreducible over  $F(\sqrt{\Delta})$

**Therefore, the possible Galois groups over  $\mathbb{Q}$  of an irreducible quartic polynomial  $f(x)$  whose discriminant is negative will be  $F(\sqrt{\Delta})$**

11. a

Splitting field of a polynomial with coefficients in a field is a smallest field extension of that field over which the polynomial splits or divides into linear factors.

Let  $F = \mathbb{Q}$  and let  $K$  be the splitting field of the polynomial  $f(x) = x^4 - 2$  over  $F$ . the roots are  $\alpha, -\alpha, i\alpha, -i\alpha$  with  $\alpha = \sqrt[4]{2}$

a.

To determine: the Galois group  $G = G(K/F)$  and the subgroup  $H = G(K/F(i))$

Since,

$$x^4 - 2 = 0$$

That is;

$$x = \sqrt[4]{2}$$

Let  $\alpha$  be some root, so;

$$\alpha = \sqrt[4]{2}$$

This implies that  $K = \mathbb{Q}(i, \sqrt[4]{2})$  is the splitting field for the polynomial

Since,

$$L = \mathbb{Q}(\sqrt[4]{2})$$

That is the total degree of the extension is;

$$2 \times 4 = 8$$

But then,

$$\text{Gal}(K/\mathbb{Q}) \leq S_4$$

This is a subgroup of  $S_4$  of order 8.

This implies that it is a Sylow-2 subgroup of, all of which are isomorphic

Also, it is known that  $D_8$ , the Dihedral group of order 8 is such a subgroup so, that it gives an isomorphism

**Therefore, the Galois group is  $G = D_8$  and the subgroup is  $H = S_4$**

b.

To show: that each element of  $H$  permutes the roots of  $f$

Now, every element of  $\text{Gal}(F/K)$  must permute the roots of  $x^4 - 2$

Further, since  $F$  is the splitting field of the polynomial that generates its roots from the subgroup  $H = G(K/F(i))$

Now, label the roots as  $\alpha, -\alpha, i\alpha, -i\alpha$  and consider;

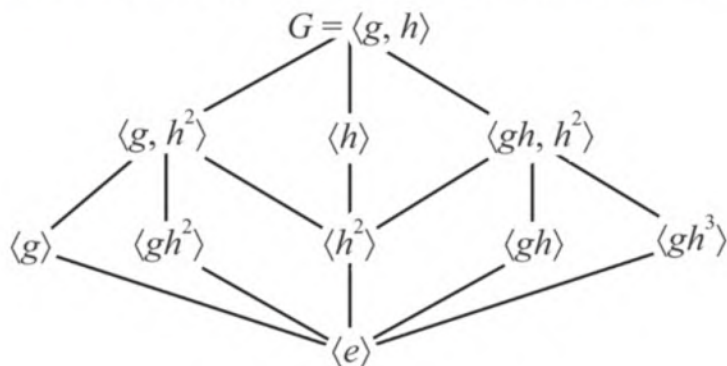
$$\text{Gal}(F/K) \subseteq S_4$$

Now, take the permutation as;

$$h = (1, 3, 2, 4)$$

And it is obvious that  $ghg = h^{-1}$  so, that  $g, h$  generates the Dihedral group of order 8.

Thus, the lattice of the subgroups that is the permutation of  $H$  to the roots of  $f$  is given below;

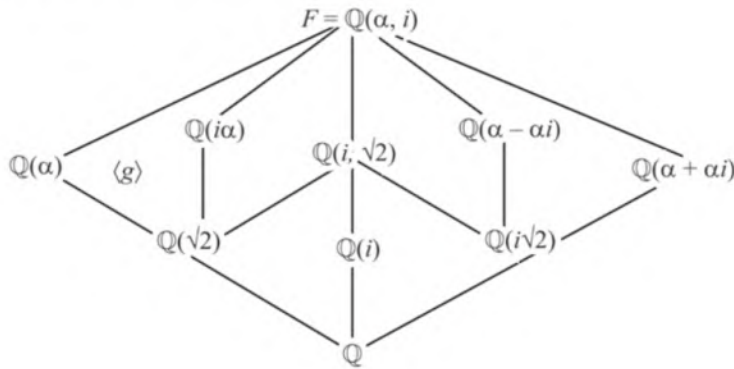


c.

To find: all intermediate fields

Now, the fixed field of the group generated by  $g, h^2$  is the intersection of the fixed fields for  $g$  and for  $h^2$  which were  $\mathbb{Q}(\alpha)$  and  $\mathbb{Q}(i, \sqrt{2})$  respectively. The intersection of these is  $\mathbb{Q}(\sqrt{2})$ . Similarly, the fixed field of group generated by  $gh, h^2$  is the intersection of  $\mathbb{Q}(\alpha - i\alpha)$  with  $\mathbb{Q}(i, \sqrt{2})$  so, the desired fixed field must be  $\mathbb{Q}(i\sqrt{2})$  and  $i\sqrt{2}$  is in both fields, so, the desired fixed field must be  $\mathbb{Q}(i\sqrt{2})$ .

The lattice that is the intermediate fields are given below;



12. a

To determine: the Galois group of each of the following expression over  $\mathbb{Q}$

a.

Consider the expression;

$$x^4 + 4x^2 + 2$$

This polynomial is irreducible by Einstein criterion

The resolvent polynomial for this polynomial will be;

$$g(x) = x^3 - 4x^2 - 8x + 32$$

This has 4 as a root

Then;

$$g(x) = (x - 4)(x^2 - 8)$$

This polynomial has only one rational root.

Hence, by Lagrange's theory;

$$\begin{aligned} G &= C_4 \\ &= D_4 \end{aligned}$$

Now, the polynomial of the form  $x^4 + ax^2 + b$  can be factored of the form as;

$$\begin{aligned} (x - \alpha)(x + \alpha)(x - \beta)(x + \beta) &= (x^2 - \alpha^2)(x^2 - \beta^2) \\ &= x^4 - (\alpha^2 + \beta^2)x^2 + \alpha^2\beta^2 \end{aligned}$$

For the polynomial  $x^4 + 4x^2 + 2$  this formula gives;

$$x^4 + 4x^2 + 2 = x^4 - (\alpha^2 + \beta^2)x^2 + \alpha^2\beta^2$$

So that;

$$\begin{aligned} \alpha^2 + \beta^2 &= -4 \\ \alpha^2\beta^2 &= 2 \end{aligned}$$

Next consider  $\alpha^2$  and  $\beta^2$ , the roots of  $x^4 + ax^2 + b$

Further by the quadratic formula,

$$\{\alpha^2, \beta^2\} = -2 \pm \sqrt{2}$$

Hence;

$$\mathcal{Q}[\alpha^2] = \mathcal{Q}[\sqrt{2}]$$

And, since;

$$\alpha^2 \beta^2 = 2$$

That is;

$$\alpha\beta = \pm\sqrt{2}$$

Thus;

$$\alpha\beta \in \mathcal{Q}[\alpha^2]$$

So,  $\mathcal{Q}[\alpha, \beta]$ , the splitting field  $L$  of  $x^4 + 4x^2 + 2$  is equals  $\mathcal{Q}[\alpha]$

That is,  $[L : K] = 4$

**Therefore, the Galois group is  $C_4$**

**b.**

Consider the expression;

$$x^4 + 2x^2 + 4$$

Now, use the result which states that;

Let  $f(X)$  be an irreducible quartic in  $\mathcal{Q}[X]$ . If  $\text{disc } f > 0$  then  $G_f = \mathbb{Z}/4\mathbb{Z}$  and therefore if  $\text{disc } f < 0$  then  $G_f = D_4$

The discriminant of the expression of the form;

$$ax^4 + cx^2 + e = 0$$

The discriminant is given by;

$$\Delta = 256a^3c^3 - 128a^2c^2e^2 + 16ac^4e$$

Comparing the values;

$$a = 1$$

$$c = 2$$

$$e = 4$$

Now, substituting the values for finding the determinant;

$$\begin{aligned} \Delta &= 256a^3c^3 - 128a^2c^2e^2 + 16ac^4e \\ &= 256(1)^3(2)^3 - 128(1)^2(2)^2(4)^2 + 16(1)(2)^4(4) \\ &= 9216 \\ &> 0 \end{aligned}$$

This means, that;

$$\text{disc } f > 0$$

**Therefore, the Galois group is  $\mathbb{Z}/4\mathbb{Z}$**

**c.**

Consider the expression;

$$x^4 + 1$$

Let  $F$  be the splitting field for;

$$f(x) = x^4 + 1$$

Then the polynomial  $f(x)$  has roots  $\{\sqrt[4]{-1}, \zeta_4 \sqrt[4]{-1}, \zeta_4^2 \sqrt[4]{-1}, \zeta_4^3 \sqrt[4]{-1}\}$

That is;

$$\begin{aligned}\sqrt[4]{-1} &= i^{\frac{1}{2}} \\ &= e^{\frac{\pi i}{4}} \\ &= \frac{\sqrt{2} + i\sqrt{2}}{2} \\ &= \frac{(1+i)\sqrt{2}}{2}\end{aligned}$$

Since,

$$\zeta_4 = i$$

A splitting field for  $f(x)$  is;

$$F = \mathbb{Q}(\sqrt{2}, i)$$

This shows that  $G_{F/\mathbb{Q}} \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

Hence, the Galois group is  $\boxed{\mathbb{Z}_2 \times \mathbb{Z}_2}$

**d.**

Consider the expression;

$$x^4 + x + 1$$

Now, use the result which states that;

Let  $f(X)$  be an irreducible quartic in  $\mathbb{Q}[X]$ . If  $\text{disc } f > 0$  then  $G_f = \mathbb{Z}/4\mathbb{Z}$  and therefore if  $\text{disc } f < 0$  then  $G_f = D_4$

The discriminant of the expression of the form;

$$ax^4 + dx + e = 0$$

The discriminant is given by;

$$\Delta = 256a^3e^3 - 27a^2d^4$$

Comparing the values;

$$a = 1$$

$$d = 1$$

$$e = 1$$

Now, substituting the values for finding the determinant;

$$\begin{aligned}\Delta &= 256a^3e^3 - 27a^2d^4 \\ &= 256(1)^3(1)^3 - 27(1)^2(1)^4 \\ &= 229 \\ &> 0\end{aligned}$$

This means, that;

$$\text{disc } f > 0$$

Therefore, the Galois group is  $\boxed{\mathbb{Z}/4\mathbb{Z}}$

e.

Consider the expression;

$$x^4 + x^3 + x^2 + x + 1$$

For finding the Galois group first substitute;

$$x = x - 1$$

That is;

$$\begin{aligned} x^4 + x^3 + x^2 + x + 1 &= (x-1)^4 + (x-1)^3 + (x+1)^2 + (x+1) + 1 \\ &= x^4 - 3x^3 + 4x^2 + 2x + 3 \end{aligned}$$

This polynomial satisfies the Einstein's criterion for the prime 5, this implies that

$$x^4 + x^3 + x^2 + x + 1 \text{ is irreducible over } \mathbb{Q}$$

Now, the roots of this function are the primitive 10<sup>th</sup> roots of unity, so, it follows that;

$$[F : \mathbb{Q}] = 4$$

That is;

$$\text{Gal}(F/\mathbb{Q}) \cong Z_{10}$$

Therefore, the Galois group is  $\boxed{Z_{10}}$

f.

Consider the expression;

$$x^4 + x^2 + 1$$

Now, use the result which states that;

Let  $f(X)$  be an irreducible quartic in  $\mathbb{Q}[X]$ . If  $\text{disc } f > 0$  then  $G_f = \mathbb{Z}/4\mathbb{Z}$  and therefore if  $\text{disc } f < 0$  then  $G_f = D_4$

The discriminant of the expression of the form;

$$ax^4 + cx^2 + e = 0$$

The discriminant is given by;

$$\Delta = 256a^3c^3 - 128a^2c^2e^2 + 16ac^4e$$

Comparing the values;

$$a = 1$$

$$c = 1$$

$$e = 1$$

Now, substituting the values for finding the determinant;

$$\begin{aligned} \Delta &= 256a^3c^3 - 128a^2c^2e^2 + 16ac^4e \\ &= 256(1)^3(1)^3 - 128(1)^2(1)^2(1)^2 + 16(1)(1)^4(1) \\ &= 144 \\ &> 0 \end{aligned}$$

This means, that;

$$\text{disc } f > 0$$

Therefore, the Galois group is  $\boxed{\mathbb{Z}/4\mathbb{Z}}$

13. a



Consider the provided statement to determine the Galois group  $G$  of  $\frac{K}{\mathbb{Q}}$  and also determine all intermediate fields. Given polynomial function is,

$$x^4 - 2x^2 - 1$$

Further, the given polynomial function can be written in the form of  $(x^2)^2 - (2x^2) - 1$  where it is compare with quadratic equation as  $a = 1, b = -2, c = -1$  then,

$$\begin{aligned} x^2 &= \frac{-(-2) \pm \sqrt{(-2)^2 - 4 \cdot 1 \cdot (-1)}}{2 \cdot 1} \\ &= \frac{2 \pm \sqrt{8}}{2} \\ &= \frac{2 \pm 2\sqrt{2}}{2} \\ &= 1 \pm \sqrt{2} \end{aligned}$$

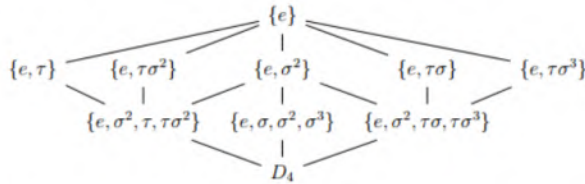
As the roots are  $1 + \sqrt{2}, 1 - \sqrt{2}$  then,  $\alpha_1 = \sqrt{1 + \sqrt{2}}, \alpha_2 = \sqrt{1 - \sqrt{2}}, \alpha_3 = -\alpha_1, \alpha_4 = -\alpha_2$

Let  $\beta_1 = 2i, \beta_2 = -2, \beta_3 = -2i$  therefore  $g(x) = (x+2)(x^2+4)$ . As in the previous exercise  $G$  is a subgroup of  $D_4 = \langle \sigma, \tau \rangle$ , where  $\sigma = (1234), \tau = (24)$  therefore by proposition 16.9.8  $G = D_4$ .

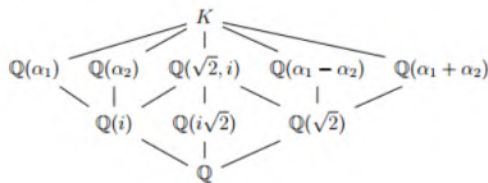
As  $\rho = \sigma^2 = (13)(24)$  then it corresponds to an automorphism  $\frac{K}{\mathbb{Q}}$ . This is called  $N$  the normal subgroup of order 2 which is generated by  $\rho$ . As  $\alpha_1^2 = 1 + \sqrt{2}$  and  $\alpha_1 \alpha_2 = i$  both are fixed by  $\rho$ . Hence,  $\mathbb{Q}(\sqrt{2}, i) \subset K^N$

Therefore from the theorem 16.5.4, the chain of fields is provided as  $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}, i) \subset K^N \subset K$  has  $[\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}] = 4, [K : K^N] = 2$ .

The lattice diagram for the subgroup of  $D_4$  is provided as below:



This lattice diagram corresponds to the little diagram of intermediate fields is provided as by computing fixed fields,



14. a

Let  $F = \mathbb{Q}(\omega)$  where  $\omega = e^{2\pi i/3}$

a.

To determine: The Galois group over  $F$  of the splitting field of  $\sqrt[3]{2+\sqrt{2}}$

Let;

$$\alpha = \sqrt[3]{2+\sqrt{2}}$$

With the splitting field  $K/F$

Here,  $\alpha$  satisfies the polynomial ;

$$\begin{aligned} f(x) &= (x^3 - 2)^2 - 2 \\ &= x^6 - 4x^3 + 2 \end{aligned}$$

The quadratic equation gives that;

$$x^3 = 2 \pm \sqrt{2}$$

Hence;

$$\alpha' = \sqrt[3]{2-\sqrt{2}}$$

Write the roots as;

$$\alpha_1 = \alpha$$

$$\alpha_2 = \alpha'$$

$$\alpha_3 = \omega\alpha$$

$$\alpha_4 = \omega\alpha'$$

$$\alpha_5 = \omega^2\alpha$$

$$\alpha_6 = \omega^2\alpha'$$

Thus, if;

$$\alpha_i \rightarrow \alpha_i$$

Then;

$$\alpha_3 \rightarrow \omega\alpha_i$$

$$\alpha_5 \rightarrow \omega^2\alpha_i$$

The permutations with the property are generated by;

$$\sigma = (123456)$$

$$\tau = (246)$$

This is in  $S_6$

These form what is called the semi direct product  $C_6 \times C_3$  and so;

$$G(K/F) \leq C_6 \times C_3$$

Now,  $f$  is irreducible over  $F$  by Eisenstein's criterion using  $p = 2$  which is prime by using the result which states that;

$$\text{Let } R = \mathbb{Z}[\omega]$$

With;

$$\omega = e^{2\pi i/3}$$

Let  $p$  be a prime integer not equals to 3 then  $(p)$  is the maximal ideal of  $R$  if and only if

$$p \equiv -1 \pmod{3}$$

Hence;

$$[F(\alpha):F] = 6$$

Also;

$$\sqrt{2} \in F(\alpha)$$

Hence,  $\alpha'$  has degree 3 or 1 over  $F(\alpha)$

Thus;

$$[K:F] = 6 \text{ or } 18$$

Claim:  $[K : F] = 18$

Consider;

$$\sigma^2 = (135)(246)$$

It is in all subgroups of  $C_6 \times C_3$  of order 6, hence, extends to an  $F$ -Automorphism of  $K$

Let  $N$  be the subgroup of order 3 generated by  $\sigma^3$

The fixed field  $K^N$  contains;

$$\alpha^2 \alpha' = \sqrt[3]{2(2 + \sqrt{2})}$$

Let  $L$  be the field generated by this element over  $F$ . The claim  $F \subset L \subset K^N \subset K$

This has by using the result of the theorem which states that;

Let  $H$  be a finite group of automorphisms of a field  $K$  and let  $F = K^H$  be its fixed field. Then  $K$  is a finite extension of  $F$ , and its degree  $[K : F]$  is equal to the order  $|H|$  of the group. Then;

$$[K : K^N] = 3$$

And;

$$[L : F] \geq 3$$

Hence;

$$[K : F] = 18$$

And so;

$$G(K/F) = C_6 \times C_3$$

b.

To determine: The Galois group over  $F$  of the splitting field of  $\sqrt{2 + \sqrt[3]{2}}$

Let;

$$\alpha = \sqrt{2 + \sqrt[3]{2}}$$

With splitting field  $K/F$

Here,  $\alpha$  satisfies the polynomial;

$$\begin{aligned} f(x) &= (x^2 - 2)^3 - 2 \\ &= x^6 - 6x^4 + 12x^2 - 10 \end{aligned}$$

The root of  $f(x)$  are;

$$\begin{aligned} \alpha_1 &\rightarrow \alpha \\ \alpha_2 &= \sqrt{2 + \omega \sqrt[3]{2}} \\ \alpha_3 &= \sqrt{2 + \omega^2 \sqrt[3]{2}} \\ \alpha_4 &= -\alpha_1 \\ \alpha_5 &= -\alpha_2 \\ \alpha_6 &= -\alpha_3 \end{aligned}$$

Thus if;

$$\alpha_i \rightarrow \alpha_j$$

Then;

$$\alpha_{i+3} \rightarrow \alpha_{j+3}$$

The group  $G$  of permutations satisfying these is generated by the subgroup of  $S_6$  isomorphic to  $S_3$  permuting  $\{\{1, 4\}, \{2, 5\}, \{3, 6\}\}$  isomorphic to  $S_3$  and;

$$\begin{aligned} C_2 &\approx \langle (14)(25)(36) \rangle \\ &\leq S_6 \end{aligned}$$

Now,  $S_3 \cap C_2 = \{1\}$  and if;

$$\sigma \in S_3$$

$$\tau \in C_2$$

$$\sigma\tau = \tau\sigma$$

Hence,  $S_3, C_2$  are normal in  $G$ ,  $G \cong S_3 \times C_2$  and  $G(K/F) \leq G$

Claim:  $[K:F] = 12$

Consider;

$$\rho = (\{1, 4\} \{2, 5\} \{3, 6\})$$

$$\in G$$

It is contained in every subgroup of  $G$  order 6, hence extends to an  $F$ -Automorphism of  $K$ .

Let  $N$  be the subgroup of degree 3 generated by  $\rho$ . The fixed field  $K^N$  contains;

$$\alpha_1\alpha_2\alpha_3 = \sqrt{10};$$

Let  $L$  be the field generated by this element over  $F$ . The chain  $F \subset L \subset K^N \subset K$  has

$$[K:K^N] = 3 \text{ by the above discussed fixed field theorem and } [L:F] = 2$$

But;

$$[K^N:L] > 1$$

Since;

$$\alpha_1 + \alpha_2 + \alpha_3 \in K^N \setminus L$$

**Hence,  $[K:F] = 12$  and so  $G(K/F) = S_3 \times C_2$**

15. a

Nested radical expression is defined as the radical expression which contains a square root sign and then contains another radical expression with a square root.

[Comment](#)

Step 2 of 3 ^

Let  $K$  be the splitting field of an irreducible quartic polynomial  $f(x)$  over  $F$  and let the roots of  $f(x)$  in  $K$  be  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$

Suppose that the resolvent cubic  $g(x)$  has a root  $\beta_1 = \alpha_1\alpha_2 + \alpha_3\alpha_4$  in  $F$

To express: The root  $\alpha_1$  explicitly in terms of nested square roots

Let  $K$  be a splitting field over  $F$  of an irreducible polynomial  $f$  of degree  $n$  in  $F[x]$  and let  $D$  be the discriminant of  $f$

The Galois group  $G(K/F)$  is a subgroup of the alternating group  $A_n$  if and only if  $D$  is a square in  $F$

By the Lagrange expressions the roots  $\alpha_i$  of quartic polynomial will be of the form;

$$\beta = \alpha_1\alpha_2 + \alpha_3\alpha_4$$

And, let;

$$g(x) = (x - \beta)$$

Now, consider the four roots of the polynomial to be  $\pm\alpha, \pm\alpha'$ , where  $\alpha$  is the nested square root

That is;

$$\alpha_1 = \alpha$$

$$\alpha_2 = -\alpha$$

$$\alpha_3 = \alpha'$$

$$\alpha_4 = -\alpha'$$

This implies that;

$$\alpha_2 = -\alpha_1$$

$$\alpha_4 = -\alpha_3$$

Now, substituting the values;

$$\beta = \alpha_1(-\alpha_1) + \alpha_3(-\alpha_3)$$

$$= -\alpha_1^2 - \alpha_3^2$$

$$\alpha_1^2 = -\beta - \alpha_3^2$$

Therefore, the root  $\alpha_1$  explicitly in terms of nested square roots is  $\boxed{\alpha_1^2 = -\beta - \alpha_3^2}$

16. a

The quartic can be solved by changing it in the general form that would allow it to be algebraically factorable and then finding the condition to put in this form. The equation that must be solved to make it factorable is called the resolvent cubic.

[Comment](#)

Step 2 of 5 ^

To determine: the resolvent cubic of the general quartic polynomial

Consider the quartic polynomial;

$$f(x) = x^4 + a_1x^3 + a_2x^2 - a_3x + a_4$$

Introduce a new variable  $t = x + \frac{a_1}{4}$

Now remove the cube term, then the equation becomes;

$$t^4 + pt^2 + qt + r = 0$$

Suppose the equation is;

$$t^4 + t^2 + 2t + 1 = 0$$

Here, the quartic part is a complete square. Therefore, the equation can be rewritten as;

$$\begin{aligned} t^4 + (t+1)^2 &= (t^2 + it + i)(t^2 - it - i) \\ &= 0 \end{aligned}$$

This is obtained by using the property;

$$a^2 + b^2 = (a + ib)(a - ib)$$

---

[Comment](#)

---

Step 4 of 5 ^

Now, consider the equation;

$$t^4 + t^2 - 4t - 3 = 0$$

Here, the quartic part is not a complete square but after shifting  $t^2$  by 1, then the following is obtained;

$$\begin{aligned} (t^2 + 1)^2 - (t^2 + 4t + 4) &= (t^2 - t - 1)(t^2 + t + 3) \\ &= 0 \end{aligned}$$

And, it is equivalent to two quartic equations

Now, the general case of the equation;

$$t^4 + pt^2 + qt + r = 0$$

With indeterminate  $\beta$  it can be written in the form;

$$(t^2 + \beta)^2 + (p - 2\beta)t^2 + qt + (r - \beta^2) = 0$$

Now, use that the quartic part is a complete square if and only if its discriminant is zero, that is;

$$\begin{aligned} D &= q^2 - 4(p - 2\beta)(r - \beta^2) \\ &= 0 \end{aligned}$$

This gives the cube equation for  $\beta$

17. a



Degree is defined as the maximum power of any given function that the largest exponential number raised to the power of the variable.

[Comment](#)

#### Step 2 of 4 ^

To determine: the real numbers  $\alpha$  of degree 4 over  $\mathbb{Q}$  that can be constructed with ruler and compass, in terms of the Galois groups of their irreducible polynomials.

For the construction let  $\alpha$  be a real number and a root of an irreducible polynomial  $f$  over the rationals.

If  $\alpha$  is constructible then,  $\mathbb{Q}(\alpha)$  is an extension of degree power of 2 but  $\mathbb{Q}(\alpha)$  is not necessarily a splitting field.

Claim: the degree of the splitting field of  $f$  is a power of 4

Now, compute a root of  $\alpha$  of  $f$  by constructing a chain of quadratic extension;

$$\mathbb{Q} < \mathbb{Q}(\alpha_1) < \mathbb{Q}(\alpha_2) < \dots < \mathbb{Q}(\alpha_n)$$

Where  $\alpha_n = \alpha$

Now, compute another root say,  $\beta$  of  $f$  by choosing the sign in some of the square roots, giving another chain of quadratic extensions;

$$\mathbb{Q} < \mathbb{Q}(\beta_1) < \mathbb{Q}(\beta_2) < \dots < \mathbb{Q}(\beta_n)$$

Where,  $\beta_n = \beta$

Further;

$$\mathbb{Q}(\beta_1) = \mathbb{Q}(\alpha_1)$$

This is not necessarily be true for the other extensions

The latter two extensions are not the same, since the first one is real and the second one is not

In this case, the Galois group is  $D_4$ , the dihedral group of order 8.

The trick is to build the splitting the field by merging the chains like

$$\mathbb{Q} < \mathbb{Q}(\alpha_1) < \mathbb{Q}(\alpha_2) < \dots < \mathbb{Q}(\alpha_n) \text{ together, to form;}$$

$$\mathbb{Q} < \mathbb{Q}(\alpha_1) < \mathbb{Q}(\alpha_1, \alpha_2) < \mathbb{Q}(\alpha_1, \alpha_2, \beta_2) < \mathbb{Q}(\alpha_1, \alpha_2, \beta_2, \alpha_3) \dots$$

This chain will terminate in the splitting field of  $f$

[Comment](#)

#### Step 4 of 4 ^

Now, it is known that  $\mathbb{Q}(\alpha_1, \beta_2)$  is of degree 2 over  $\mathbb{Q}(\alpha_1)$

This means that  $\mathbb{Q}(\alpha_1, \alpha_2, \beta_2)$  is of degree 1 or 2 over  $\mathbb{Q}(\alpha_1, \alpha_2)$ , since  $\beta_2$  satisfies a quadratic equation over  $\mathbb{Q}(\alpha_1)$  and that equation is also quadratic equation over  $\mathbb{Q}(\alpha_1, \alpha_2)$

**This implies that the value of  $\alpha$  will depend upon the degree of the irreducible function  $f$ , that is it will be mostly of the form of fourth degree of an equation**

18. a

The Klein four-group is defined as the group  $\mathbb{Z}_2 \times \mathbb{Z}_2$  that is the direct product of two copies of the cyclic group of order 2.

To prove that any Galois extension whose Galois group is the dihedral group  $D_4$  is the splitting field of a polynomial of the form  $x^4 + bx^2 + c$

On the contrary assume that the polynomial  $x^4 + bx^2 + c \in \mathbb{Q}[x]$  is irreducible

Claim; that its Galois group is the Klein subgroup if  $\sqrt{b} \in \mathbb{Q}$ , the cyclic group of order 4 if;

$$\sqrt{a^2 - 4b}\sqrt{2} \in \mathbb{Q}$$

If this is not true than any Galois extension whose Galois group is the dihedral group  $D_4$  is the splitting field of a polynomial of the form  $x^4 + bx^2 + c$

It is already known that the possible Galois groups are  $K_4, \mathbb{Z}_4$  or  $D_4$

The roots are  $\alpha, \beta, -\alpha, -\beta$  which satisfy the following relations;

$$\alpha\beta = \sqrt{b}$$

$$\alpha^2 - \beta^2 = \sqrt{a^2 - 4b}$$

$$\alpha^3\beta - \beta\alpha^3 = \sqrt{b}\sqrt{a^2 - 4b}$$

If  $\sqrt{b} \in \mathbb{Q}$

Then;

$$\alpha\beta \in \mathbb{Q}$$

Let  $s \in G$  be such that;

$$s(\alpha) = \beta$$

Then;

$$\begin{aligned} s(\beta) &= \frac{\sqrt{b}}{s(\alpha)} \\ &= \alpha \end{aligned}$$

Similarly, if;

$$s(\alpha) = -\beta$$

$$s(-\beta) = \alpha$$

Finally if;

$$s(\alpha) = -\alpha$$

$$s(\beta) = -\beta$$

Thus, every element of Galois group has order 2. this implies that the Galois group is the Klein group

Now, assume that;

$$\sqrt{b}\sqrt{a^2 - 4b} \in \mathbb{Q}$$

Then;

$$\alpha^3\beta - \beta\alpha^3 \in \mathbb{Q}$$

Let  $s$  be an element of the Galois groups which maps  $\alpha$  to  $\beta$

$$\text{If } s(\beta) = \alpha$$

Then;

$$s(\alpha^3\beta - \beta^3\alpha) = \beta^3\alpha - \alpha\beta^3$$

This is impossible. Therefore;

$$s(\beta) = -\alpha$$

Thus,  $s$  must have order 4, implies that the Galois group is  $\mathbb{Z}_4$

Finally, the splitting field must contain  $\mathbb{Q}(\sqrt{b}), \mathbb{Q}(\sqrt{a^2-b})$  and  $\mathbb{Q}(\sqrt{b}\sqrt{a^2-b}-4b)$

[Comment](#)

Step 5 of 5 ^

The irreducibility of the polynomial implies that  $\sqrt{a^2-4b}$  is not rational. Therefore, if;

$$\sqrt{b}, \sqrt{a^2-4b}\sqrt{b} \notin \mathbb{Q}$$

The splitting field contains at least three subfields of degree 2.

Hence, the Galois group is either  $K_4$  or  $D_4$

However, if the group is  $K_4$ , then  $\alpha\beta$  is fixed by any element of the Galois group.

Since,  $\sqrt{b}$  is not rational.

Thus, the only possibility is  $D_4$

Therefore; any Galois extension whose Galois group is the dihedral group  $D_4$  is the splitting field of a polynomial of the form  $x^4 + bx^2 + c$

## Section 10

1. a

**Solution:** We will determine the degree of  $\zeta_7$  over the field  $\mathbb{Q}(\zeta_3)$ .

Now notice that the degree of  $\zeta_3$  over  $\mathbb{Q}$  is 2.

Therefore its degree over  $\mathbb{Q}(\zeta_7)$  is either 2 or 1. Let us consider that the degree is 1.

Then notice that

$$\zeta_3 \in \mathbb{Q}(\zeta_7).$$

Therefore  $\mathbb{Q}(\zeta_7)$  contains the subfield  $\mathbb{Q}(\sqrt{-3})$ .

Now recall that, let  $p$  be a prime different from 2, and let  $L$  be the unique quadratic extension of  $\mathbb{Q}$  contained in the cyclotomic field  $\mathbb{Q}(\zeta_p)$ . Now if  $p \equiv 1 \pmod{4}$ , then  $L = \mathbb{Q}(\sqrt{p})$  and if  $p \equiv 3 \pmod{4}$ , then  $L = \mathbb{Q}(\sqrt{-p})$ .

It follows that the unique quadratic extension of  $\mathbb{Q}$  contained in  $\mathbb{Q}(\zeta_7)$  is  $\mathbb{Q}(\sqrt{-7})$ .

Therefore we contradict the above fact. Therefore we have

$$[\mathbb{Q}(\zeta_7, \zeta_3) : \mathbb{Q}(\zeta_7)] = 2.$$

And now note that

$$[\mathbb{Q}(\zeta_7, \zeta_3) : \mathbb{Q}(\zeta_3)] = \frac{[\mathbb{Q}(\zeta_7, \zeta_3) : \mathbb{Q}(\zeta_7)][\mathbb{Q}(\zeta_7) : \mathbb{Q}]}{[\mathbb{Q}(\zeta_3) : \mathbb{Q}]} = 6.$$

Thus the degree of  $\zeta_7$  over the field  $\mathbb{Q}(\zeta_3)$  is 6.

This completes the solution.

## Result

The degree of  $\zeta_7$  over the field  $\mathbb{Q}(\zeta_3)$  is 6.

## 2. a

Generator of any set is defined as the set of group of elements which are contained in the whole but not in any of the subgroups.

[Comment](#)

Step 2 of 3 ^

First consider;

$$\zeta = \zeta_{17}$$

Clearly, the residue of 3 is a primitive root modulo 17, so the Galois group  $G = G(K/F)$  is a cyclic group of order 16.

Since;

$$\zeta \rightarrow \zeta^3$$

The fixed fields of the subgroups can be defined as;

$$\begin{aligned} F &= L_0 \\ &= K^{(\sigma^0)} \end{aligned}$$

And;

$$\begin{aligned} L_1 &= K^{(\sigma^2)} \\ L_2 &= K^{(\sigma^4)} \end{aligned}$$

Let  $\beta_1, \beta_2$  denote the orbit sums. Then  $\{\beta_1, \beta_2\}$  is a  $G$ -orbit. So, the irreducible polynomial for  $\beta_1$  and  $\beta_2$  is;

$$(x - \beta_1)(x - \beta_2)$$

To further solve the polynomial, consider as follows;

$$p_1(\beta) = \beta_1 + \beta_2$$

$$p_2(\beta) = \beta_1\beta_2$$

Since,  $p_1$  is the sum of all  $\zeta^i$  this means that;

$$p_1(\beta) = -1$$

Now,  $p_2(\beta) \in \mathbb{Q}$  so, expanding the term  $\beta_1, \beta_2$  gives 64 summands of the form of  $\zeta^i$  appears four times, thus;

$$p_2(\beta) = -4$$

Then, the polynomial becomes;

$$(x - \beta_1)(x - \beta_2) = x^2 + x - 4$$

Also, its discriminant is 17. So;

$$K^{(\sigma^3)} = F(\sqrt{17})$$

Similarly;

$$K^{(\sigma^4)} = F(\sqrt[4]{17})$$

**Therefore, the generators of  $\zeta = \zeta_{17}$  for the intermediate field  $L_2$  is  $F(\sqrt[4]{17})$**

### 3. a

Consider the statement to find the degree of the given element over  $\mathbb{Q}$  with  $\zeta = \zeta_7$ .

If it is assume that to find the degree of  $\alpha \in \mathbb{Q}(\zeta)$  then from proposition 16.10.2,

$$G\left(\frac{\mathbb{Q}(\zeta)}{\mathbb{Q}}\right) = C_6$$

It consists of  $\sigma_i$  which is given as  $\sigma_i(\zeta) = \zeta^i$  where  $1 \leq i \leq 6$ . If  $H \leq C_6$  is the subgroup fixing  $\alpha$  then from the theorem 16.7.1,

$$\mathbb{Q}(\alpha) = \mathbb{Q}(\zeta)^H$$

Therefore, from the multiplicative property of degree it gives the result as below:

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{6}{|H|} \end{aligned}$$

(a)

Let  $\zeta = \zeta_7$  and the provided element is,  $\zeta + \zeta^5$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{6}{|H|} \end{aligned}$$

So  $\{\sigma_1\}$  fixes  $\zeta + \zeta^5$  therefore the degree of  $\zeta + \zeta^5$  is  $\boxed{6}$ .

(b)

Let  $\zeta = \zeta_7$  and the provided element is,  $\zeta^3 + \zeta^4$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)'']} \\ &= \frac{6}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_6\}$  fixes  $\zeta^3 + \zeta^4$  therefore the degree of  $\zeta^3 + \zeta^4$  is  $\boxed{3}$ .

(c)

Let  $\zeta = \zeta_7$  and the provided element is,  $\zeta^3 + \zeta^5 + \zeta^6$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)'']} \\ &= \frac{6}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_2, \sigma_4\}$  fixes  $\zeta^3 + \zeta^5 + \zeta^6$  therefore the degree of  $\zeta^3 + \zeta^5 + \zeta^6$  is  $\boxed{2}$ .

4. a

Consider the statement to find the degree of the given element over  $\mathbb{Q}$  with  $\zeta = \zeta_{13}$ .

If it is assume that to find the degree of  $\alpha \in \mathbb{Q}(\zeta)$  then from proposition 16.10.2,

$$G\left(\frac{\mathbb{Q}(\zeta)}{\mathbb{Q}}\right) = C_{12}$$

It consists of  $\sigma_i$  which is given as  $\sigma_i(\zeta) = \zeta^i$  where  $1 \leq i \leq 12$ . If  $H \leq C_{12}$  is the subgroup fixing  $\alpha$  then from the theorem 16.7.1,

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] = [\mathbb{Q}(\zeta) : \mathbb{Q}] / |H|$$

Therefore, from the multiplicative property of degree it gives the result as below:

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)'']} \\ &= \frac{12}{|H|} \end{aligned}$$

(a)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta + \zeta^{12}$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)'']} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_2\}$  fixes  $\zeta + \zeta^{12}$  therefore the degree of  $\zeta + \zeta^{12}$  is  $\boxed{6}$ .



(b)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta + \zeta^2$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_{12}\}$  fixes  $\zeta + \zeta^2$  therefore the degree of  $\zeta + \zeta^2$  is  $\boxed{6}$ .

(c)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta + \zeta^5 + \zeta^8$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_9, \sigma_6\}$  fixes  $\zeta + \zeta^5 + \zeta^8$  therefore the degree of  $\zeta + \zeta^5 + \zeta^8$  is  $\boxed{4}$ .

(d)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta^2 + \zeta^5 + \zeta^6$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_{12}, \sigma_9, \sigma_8\}$  fixes  $\zeta^2 + \zeta^5 + \zeta^6$  therefore the degree of  $\zeta^2 + \zeta^5 + \zeta^6$  is  $\boxed{4}$ .

(e)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta + \zeta^5 + \zeta^8 + \zeta^{12}$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_9, \sigma_6, \sigma_2\}$  fixes  $\zeta + \zeta^5 + \zeta^8 + \zeta^{12}$  therefore the degree of  $\zeta + \zeta^5 + \zeta^8 + \zeta^{12}$  is  $\boxed{3}$ .

(f)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta + \zeta^2 + \zeta^5 + \zeta^{12}$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_{12}, \sigma_9, \sigma_2\}$  fixes  $\zeta + \zeta^2 + \zeta^5 + \zeta^{12}$  therefore the degree of  $\zeta + \zeta^2 + \zeta^5 + \zeta^{12}$  is  $\boxed{3}$ .

(g)

Let  $\zeta = \zeta_{13}$  and the provided element is,  $\zeta + \zeta^3 + \zeta^4 + \zeta^9 + \zeta^{10} + \zeta^{12}$

From the above calculation it is observed that, the degree is provided as below from the multiplicative property,

$$\begin{aligned} [\mathbb{Q}(\alpha) : \mathbb{Q}] &= \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[\mathbb{Q}(\zeta) : \mathbb{Q}(\zeta)^H]} \\ &= \frac{12}{|H|} \end{aligned}$$

So  $\{\sigma_1, \sigma_{11}, \sigma_{10}, \sigma_3, \sigma_4, \sigma_2\}$  fixes  $\zeta + \zeta^3 + \zeta^4 + \zeta^9 + \zeta^{10} + \zeta^{12}$  therefore the degree of  $\zeta + \zeta^3 + \zeta^4 + \zeta^9 + \zeta^{10} + \zeta^{12}$  is  $\boxed{2}$ .

5. a

If  $L$  is an extension of  $F$  and  $F$  is an extension of  $K$ , that is  $L \subset F \subset K$ , then is defined as the intermediate field of the field extension  $L/K$

Let  $K = \mathbb{Q}(\zeta_p)$

a.

Consider the case for  $p = 5$

That is;

$$K = \mathbb{Q}(\zeta_5)$$

Now,  $\zeta_5$  is defined as the fifth root of unity

And, by using the Einstein criteria which states that;

$$|\mathbb{Q}(\zeta) : \mathbb{Q}| = p - 1$$

Here,  $p = 5$  so;

$$\begin{aligned} |\mathbb{Q}(\zeta) : \mathbb{Q}| &= 5 - 1 \\ &= 4 \end{aligned}$$

So, this means that any intermediate field between  $\mathbb{Q}(\zeta_5)$  and  $\mathbb{Q}$  of the form of;

$$\zeta_5^4 + \zeta_5^3 + \zeta_5^2 + \zeta_5 + 1 = 0$$

Dividing both sides by  $\zeta_5^2$ ;

$$\frac{\zeta_5^4 + \zeta_5^3 + \zeta_5^2 + \zeta_5 + 1}{\zeta_5^2} = \frac{0}{\zeta_5^2}$$

$$\zeta_5^2 + \zeta_5 + 1 + \zeta_5^{-1} + \zeta_5^{-2} = 0$$

That is;

$$\left(\zeta_5 + \frac{1}{\zeta_5}\right)^2 + \left(\zeta_5 + \frac{1}{\zeta_5}\right) - 1 = 0$$

Now let;

$$\alpha = \left(\zeta_5 + \frac{1}{\zeta_5}\right)$$

That is;

$$\begin{aligned} \alpha &= \zeta_5 + \frac{1}{\zeta_5} \\ &= e^{2\pi i/5} + e^{-2\pi i/5} \\ &= 2 \cos \frac{2\pi}{5} \end{aligned}$$

This implies that;

$$\mathbb{Q}(\sqrt{5}) \subset \mathbb{Q}$$

Hence, the intermediate field for  $K = \mathbb{Q}(\zeta_5)$  is  $\boxed{\mathbb{Q}(\sqrt{5})}$

b.

Consider the case for  $p = 7$

That is;

$$K = \mathbb{Q}(\zeta_7)$$

Now,  $\zeta_7$  is defined as the fifth root of unity

And, by using the Einstein criteria which states that;

$$[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1$$

Here,  $p = 7$  so;

$$[\mathbb{Q}(\zeta) : \mathbb{Q}] = 7 - 1 \\ = 6$$

So, this means that any intermediate field between  $\mathbb{Q}(\zeta_7)$  and  $\mathbb{Q}$  of the form of;

$$\zeta_7^6 + \zeta_7^5 + \zeta_7^4 + \zeta_7^3 + \zeta_7^2 + \zeta_7 + 1 = 0$$

Dividing both sides by  $\zeta_7^3$ ;

$$\frac{\zeta_7^6 + \zeta_7^5 + \zeta_7^4 + \zeta_7^3 + \zeta_7^2 + \zeta_7 + 1}{\zeta_7^3} = \frac{0}{\zeta_7^3}$$

$$\zeta_7^3 + \zeta_7^2 + \zeta_7 + 1 + \zeta_7^{-1} + \zeta_7^{-2} + \zeta_7^{-3} = 0$$

That is;

$$\left(\zeta_7 + \frac{1}{\zeta_7}\right)^3 + \left(\zeta_7 + \frac{1}{\zeta_7}\right)^2 - 2\left(\zeta_7 + \frac{1}{\zeta_7}\right) - 1 = 0$$

Now let;

$$\alpha = \left(\zeta_7 + \frac{1}{\zeta_7}\right)$$

That is;

$$\alpha = \zeta_7 + \frac{1}{\zeta_7} \\ = e^{2\pi i/7} + e^{-2\pi i/7} \\ = 2 \cos \frac{2\pi}{7}$$

This implies that;

$$\mathbb{Q}(\sqrt{7}) \subset \mathbb{Q}$$

[Comment](#)

Step 6 of 10 ^

Hence, the intermediate field for  $K = \mathbb{Q}(\zeta_7)$  is  $\boxed{\mathbb{Q}(\sqrt{7})}$

c.

Consider the case for  $p = 11$

That is;

$$K = \mathbb{Q}(\zeta_{11})$$

Now,  $\zeta_{11}$  is defined as the fifth root of unity

And, by using the Einstein criteria which states that;

$$|\mathbb{Q}(\zeta)| : \mathbb{Q} = p - 1$$

Here,  $p = 11$  so;

$$|\mathbb{Q}(\zeta)| : \mathbb{Q} = 11 - 1 \\ = 10$$

So, this means that any intermediate field between  $\mathbb{Q}(\zeta_{11})$  and  $\mathbb{Q}$  of the form of;

$$\zeta_{11}^{10} + \zeta_{11}^9 + \zeta_{11}^8 + \zeta_{11}^7 + \zeta_{11}^6 + \zeta_{11}^5 + \zeta_{11}^4 + \zeta_{11}^3 + \zeta_{11}^2 + \zeta_{11} + 1 = 0$$

Dividing both sides by  $\zeta_{11}^5$ ;

$$\frac{\zeta_{11}^{10} + \zeta_{11}^9 + \zeta_{11}^8 + \zeta_{11}^7 + \zeta_{11}^6 + \zeta_{11}^5 + \zeta_{11}^4 + \zeta_{11}^3 + \zeta_{11}^2 + \zeta_{11} + 1}{\zeta_{11}^5} = \frac{0}{\zeta_{11}^5}$$

$$\zeta_{11}^5 + \zeta_{11}^4 + \zeta_{11}^3 + \zeta_{11}^2 + \zeta_{11} + 1 + \zeta_{11}^{-1} + \zeta_{11}^{-2} + \zeta_{11}^{-3} + \zeta_{11}^{-4} + \zeta_{11}^{-5} = 0$$

That is;

$$\left(\zeta_{11} + \frac{1}{\zeta_{11}}\right)^5 + \left(\zeta_{11} + \frac{1}{\zeta_{11}}\right)^4 + \left(\zeta_{11} + \frac{1}{\zeta_{11}}\right)^3 + \left(\zeta_{11} + \frac{1}{\zeta_{11}}\right)^2 - 2\left(\zeta_{11} + \frac{1}{\zeta_{11}}\right) - 1 = 0$$

Now let;

$$\alpha = \left(\zeta_{11} + \frac{1}{\zeta_{11}}\right)$$

That is;

$$\alpha = \zeta_{11} + \frac{1}{\zeta_{11}} \\ = e^{2\pi i/11} + e^{-2\pi i/11} \\ = 2 \cos \frac{2\pi}{11}$$

This implies that;

$$\mathbb{Q}(\sqrt{11}) \subset \mathbb{Q}$$

Hence, the intermediate field for  $K = \mathbb{Q}(\zeta_{11})$  is  $\boxed{\mathbb{Q}(\sqrt{11})}$

d.

Consider the case for  $p = 13$

That is;

$$K = \mathbb{Q}(\zeta_{13})$$

Now,  $\zeta_{13}$  is defined as the fifth root of unity

And, by using the Einstein criteria which states that;

$$|\mathbb{Q}(\zeta): \mathbb{Q}| = p - 1$$

Here,  $p = 13$  so;

$$|\mathbb{Q}(\zeta): \mathbb{Q}| = 13 - 1 \\ = 12$$

So, this means that any intermediate field between  $\mathbb{Q}(\zeta_{13})$  and  $\mathbb{Q}$  of the form of;

$$\zeta_{13}^{12} + \zeta_{13}^{11} + \zeta_{13}^{10} + \zeta_{13}^9 + \zeta_{13}^8 + \zeta_{13}^7 + \zeta_{13}^6 + \zeta_{13}^5 + \zeta_{13}^4 + \zeta_{13}^3 + \zeta_{13}^2 + \zeta_{13} + 1 = 0$$

Dividing both sides by  $\zeta_{13}^6$ ;

$$\frac{\zeta_{13}^{12} + \zeta_{13}^{11} + \zeta_{13}^{10} + \zeta_{13}^9 + \zeta_{13}^8 + \zeta_{13}^7 + \zeta_{13}^6 + \zeta_{13}^5 + \zeta_{13}^4 + \zeta_{13}^3 + \zeta_{13}^2 + \zeta_{13} + 1}{\zeta_{13}^6} = \frac{0}{\zeta_{13}^6}$$

$$\zeta_{13}^6 + \zeta_{13}^5 + \zeta_{13}^4 + \zeta_{13}^3 + \zeta_{13}^2 + \zeta_{13} + 1 + \zeta_{13}^{-1} + \zeta_{13}^{-2} + \zeta_{13}^{-3} + \zeta_{13}^{-4} + \zeta_{13}^{-5} + \zeta_{13}^{-6} = 0$$

That is;

$$\left(\zeta_{13} + \frac{1}{\zeta_{13}}\right)^6 + \left(\zeta_{13} + \frac{1}{\zeta_{13}}\right)^5 + \left(\zeta_{13} + \frac{1}{\zeta_{13}}\right)^4 + \left(\zeta_{13} + \frac{1}{\zeta_{13}}\right)^3 + \left(\zeta_{13} + \frac{1}{\zeta_{13}}\right)^2 - 2\left(\zeta_{13} + \frac{1}{\zeta_{13}}\right) - 1 = 0$$

Now let;

$$\alpha = \left(\zeta_{13} + \frac{1}{\zeta_{13}}\right)$$

That is;

$$\alpha = \zeta_{13} + \frac{1}{\zeta_{13}} \\ = e^{2\pi i/13} + e^{-2\pi i/13} \\ = 2 \cos \frac{2\pi}{13}$$

This implies that;

$$\mathbb{Q}(\sqrt{13}) \subset \mathbb{Q}$$

Hence, the intermediate field for  $K = \mathbb{Q}(\zeta_{13})$  is  $\boxed{\mathbb{Q}(\sqrt{13})}$

6. a

(a)

Consider the provided statement to prove the field  $\mathbb{Q}(\zeta_p)$  contains a unique quadratic extension of  $\mathbb{Q}$  such that  $\mathbb{Q}(\sqrt{\text{disc}(\phi_p(x))}) = \mathbb{Q}(\sqrt{(-1)^{\frac{p-1}{2}} p})$  that is real if  $p \equiv 1 \pmod{4}$  and complex if  $p \equiv 3 \pmod{4}$ .

[Comment](#)

Step 2 of 4 ^

The Galois group  $G$  of  $\mathbb{Q}(\zeta_p)$  over  $\mathbb{Q}$  is cyclic and which has order  $p-1$ . Therefore there is a unique subgroup of  $G$  that having index 2. So, there is a unique subfield of  $\mathbb{Q}(\zeta_p)$  that is a quadratic extension of  $\mathbb{Q}$  as  $\mathbb{Q}(\sqrt{\text{disc}(\phi_p(x))}) \in \frac{\mathbb{Q}(\zeta_p)}{\mathbb{Q}}$  that generates the unique quadratic subfield of  $\mathbb{Q}(\zeta_p)$ .

(b)

Consider the provided statement to prove the Kronecker-Weber Theorem for quadratic extensions.

As from the Kronecker-Weber Theorem, every Galois extension of the field  $\mathbb{Q}$  whose galois group is abelian is in one of the cyclotomic fields  $\mathbb{Q}(\zeta_n)$ .

[Comment](#)

Step 4 of 4 ^

If  $p \equiv 3 \pmod{4}$  then  $\mathbb{Q}(\sqrt{-p}) \subseteq \mathbb{Q}(\zeta_p)$  and if  $p \equiv 1 \pmod{4}$  then  $\mathbb{Q}(\sqrt{p}) \subseteq \mathbb{Q}(\zeta_p)$

A quadratic extension of  $\mathbb{Q}$  is of the form  $\mathbb{Q}(\sqrt{d})$  where  $d$  belongs to a square free integer.

Then, it is assumed that  $d = \pm p_1 p_2 \cdots p_r$  where  $p_1, p_2, \dots, p_r$  are distinct primes.

Therefore,  $\mathbb{Q}(\sqrt{d}) \subseteq \mathbb{Q}(\zeta_{p_1}, \zeta_{p_2}, \dots, \zeta_{p_r}, i)$  is proved.

7. a



Let  $F$  be a field and let  $K$  be an extension of the field  $F$ . Then  $K$  is said to be Galois extension of the field  $F$  if the extension  $K$  is a splitting field of some separable polynomial,

That is it is sufficient to find a polynomial over the field  $F$  whose roots lie entirely in the extension  $K$ .

[Comment](#)

Step 2 of 5 ^

(a)

Let  $\zeta_n = e^{2\pi i/n}$  and  $K = \mathbb{Q}(\zeta_n)$

Consider the polynomial  $p(x) = x^n - 1$

Now factorize the polynomial  $p(x)$

$$\begin{aligned} p(x) &= x^n - 1 \\ &= (x-1)(1+x+x^2+\dots+x^{n-1}) \end{aligned}$$

Now the expression  $(1+x+x^2+\dots+x^{n-1})$  can be factored in the following way

$$(1+x+x^2+\dots+x^{n-1}) = ((x-\zeta)(x-\zeta^2)\dots(x-\zeta^{n-1})), \text{ where } \zeta^i \text{ denotes the } i^{\text{th}} \text{ root of unity.}$$

Then the polynomial  $p(x)$  can be written as

$$p(x) = (x-1)(x-\zeta)(x-\zeta^2)\dots(x-\zeta^{n-1})$$

Thus, the above extension field is a splitting field for the separable polynomial over  $\mathbb{Q}$ .

**Hence the extension field  $K$  is a Galois extension of  $\mathbb{Q}$ .**

(b)

Let  $\sigma \in G(K/\mathbb{Q})$

Define a map  $f$  from  $G(K/\mathbb{Q})$  to  $U$ , where  $U$  is the group of units in the ring  $\mathbb{Z}/(n)$  as

$$f: G(K/\mathbb{Q}) \rightarrow U$$

Such that  $f(\sigma) = a_\sigma \bmod n$ , where  $\sigma(\zeta) = \zeta^{a_\sigma}$  for every root of unity

Let  $\sigma, \rho \in G(K/\mathbb{Q})$

Let  $\zeta_n$  be the  $n$ -th root of unity.

Consider  $\sigma\rho(\zeta_n)$

$$\begin{aligned} \sigma\rho(\zeta_n) &= \sigma(\rho(\zeta_n)) \\ &= \sigma(\zeta_n^{a_\rho}) \\ &= (\zeta_n^{a_\rho})^{a_\sigma} \\ &= \zeta_n^{a_\rho a_\sigma} \end{aligned}$$

Thus,  $\sigma\rho(\zeta_n) = \zeta_n^{a_\rho a_\sigma}$

But,  $\sigma\rho(\zeta_n) = \zeta_n^{a_{\sigma\rho}}$

Since the left hand side of above two equations are equal so right hand side must be equal too.

That is,  $\zeta_n^{a_\rho a_\sigma} = \zeta_n^{a_{\sigma\rho}}$

Now since the order of the  $n^{\text{th}}$  primitive root of unity  $\zeta_n$  is  $n$ , thus

$$a_{\sigma\rho} \equiv a_\sigma a_\rho \bmod n.$$

So, the above map is a homomorphism.

### Injectivity

Let  $\sigma \in \text{Ker}(f)$ .

This implies that  $\sigma$  gets mapped to the identity in  $U$ .

Thus,

$$a_\sigma \equiv 1 \pmod{n}$$

Consider  $\sigma(\zeta_n)$

$$\begin{aligned}\sigma(\zeta_n) &= \zeta_n^{a_\sigma} \\ &= \zeta_n\end{aligned}$$

So for every  $x \in K$ ,  $\sigma(x) = x$

Hence,  $\sigma$  is the identity element in  $K$ .

Thus  $\sigma$  is the identity in the group  $G(K/\mathbb{Q})$ .

This implies that the map is injective.

**Therefore, the result stated in the question has been proved.**

### (c)

Let  $\zeta_n$  be primitive  $n^{\text{th}}$  root of unity.

Let  $a \in U$

Then,  $a \in \mathbb{Z}$  such that  $(a, n) = 1$

In order to prove surjectivity, it suffices to show that  $\zeta_n$  and  $\zeta_n^a$  have same minimal polynomial.

Consider the prime factorization of  $a$

$$a = p_1 p_2 \dots p_r, \text{ where no } p_i \text{ divides } n$$

Instead of proving the above said statement for  $\zeta_n$  and  $\zeta_n^a$ , it is sufficient to prove this condition for  $\zeta_n$  and  $\zeta_n^{p^r}$ , where  $p$  is a prime not dividing  $n$ .

Assume on contrary that  $\zeta_n$  and  $\zeta_n^{p^r}$  do not have same minimal polynomial

Let  $f(t)$  denote the minimal polynomial for  $\zeta_n$

Let  $g(t)$  denote the minimal polynomial for  $\zeta_n^{p^r}$

Then by the assumption,  $f(t) \neq g(t)$

Since all the  $n^{\text{th}}$  roots of unity satisfy the relation  $t^n - 1 = 0$

So,  $g(t) \mid t^n - 1$  and  $f(t) \mid t^n - 1$ , and since  $f(t) \neq g(t)$

Thus,  $t^n - 1 = g(t)f(t)h(t)$ , where  $h(t) \in \mathbb{Q}[t]$

Under  $\text{mod } p$  arithmetic

$$\begin{aligned}(t^n - 1) \pmod{p} &= (g(t)f(t)h(t)) \pmod{p} \\ t^n - 1^p &= g^p(t)f^p(t)h^p(t)\end{aligned}$$

Since  $p$  does not divide  $n$  so the above polynomial is separable.

The expression  $t^n - 1^p = g^p(t)f^p(t)h^p(t)$  implies that  $f^p$  and  $g^p$  are relatively prime

Also  $f, g$  are monic polynomials so following equation holds

$$\begin{aligned}\deg(f) &= \deg(f^p) \\ \deg(g) &= \deg(g^p)\end{aligned}$$

Thus,  $f^p$  and  $g^p$  are both non-constants

Now since  $g(\zeta_n^{p^r}) = 0$

This implies that  $\zeta_n^{p^r}$  is a zero of  $g(t)$

This further implies that  $\zeta_n$  is a zero of the polynomial  $g(t^p)$

But  $\zeta_n$  is also a zero of  $f(t)$ , so  $f(t)$  divides  $g(t^p)$

Thus,  $g(t^p) = f(t)s(t)$ , where  $s(t) \in \mathbb{Q}[t]$

Under mod  $p$  arithmetic the above equation becomes

$$g^a(t)^p = f^a(t)s^a(t)$$

This implies that every irreducible factor of  $f^a(t)$  is a factor of  $g^a(t)$

This contradicts the above derived fact that  $f^a$  and  $g^a$  are relatively prime.

This contradiction arises due to the wrong assumption that  $\zeta_n$  and  $\zeta_n^p$  do not have same minimal polynomial, where  $p$  is a prime that does not divide  $n$ .

Hence  $\zeta_n$  and  $\zeta_n^p$  have same minimal polynomial.

Thus, the map defined in part (a) is surjective and thus a bijective homomorphism for every  $n$ .

**Therefore the map defined in part (a) is a bijective homomorphism for every  $n$  and in particular for  $n = 6, 8, 12$ .**

## 8. a

Galois Theory is defined as the method of applying the group theory to the given solution of the algebraic equations.

The conditions for finding the Galois group is given in the below mentioned table;

	$D$ is a square	$D$ is not a square
reducible	$D_2$	$D_4$ or $C_4$
irreducible	$A_4$	$S_4$

Where,  $D$  is the discriminant

**a.**

Consider the given polynomial;

$$x^8 - 1$$

The discriminant for this polynomial is given by;

$$\begin{aligned} x^8 - 1 &= (8)^8 (1)^7 (-1)^7 \\ &= -16777216 \end{aligned}$$

Since, the discriminant is negative so, it is not a square.

Also, the roots of  $x^8 - 1 = 0$  are  $-1, 1, -i, i, -\sqrt[4]{-1}, \sqrt[4]{-1}, -(-1)^{3/4}$

There are complex roots. So, the polynomial is irreducible in  $\mathbb{Q}$

Hence, by using the above given table for the Galois group

**The Galois group for the polynomial  $x^8 - 1$  is  $G = S_4$**

**b.**

Consider the given polynomial;

$$x^{12} - 1$$

The discriminant for this polynomial is given by;

$$\begin{aligned} x^{12} - 1 &= (12)^{12} (1)^{11} (-1)^{11} \\ &= -8916100448256 \end{aligned}$$

Since, the discriminant is negative so, it is not a square.

Also, the roots of  $x^{12} - 1 = 0$  are  $-1, 1, -i, i, -\sqrt[6]{-1}, \sqrt[6]{-1}, -(-1)^{1/3}$

There are complex roots. So, the polynomial is irreducible in  $\mathbb{Q}$

Hence, by using the above given table for the Galois group

**The Galois group for the polynomial  $x^{12} - 1$  is  $G = S_4$**

c.

Consider the given polynomial;

$$x^9 - 1$$

The discriminant for this polynomial is given by;

$$\begin{aligned} x^9 - 1 &= (9)^9 (1)^{10} (-1)^{10} \\ &= 387420489 \end{aligned}$$

Since, the discriminant is positive and also, it is a perfect square.

Also, the roots of  $x^9 - 1 = 0$  are  $-\sqrt[9]{-1}, (-1)^{2/9}, -\sqrt[9]{-1}, (-1)^{4/9}, -(-1)^{5/9}$

There are no complex roots. So, the polynomial is reducible in  $\mathbb{Q}$

Hence, by using the above given table for the Galois group

**The Galois group for the polynomial  $x^9 - 1$  is  $G = D_2$**

9. a

For a polynomial  $P(x)$  of degree  $n$  with roots  $u_1, u_2, \dots, u_n$  which is given as follows

$$P(x) = x^n - s_1 x^{n-1} + s_2 x^{n-2} - \dots \pm s_n,$$

Where  $s_i$  denote the elementary symmetric function.

For such polynomial function the discriminant is denoted by  $D(u)$  and is defined as

$$D(u) = \prod_{i < j} (u_i - u_j)^2$$

Let  $\zeta_n$  denote the primitive  $n^{\text{th}}$  root of unity over the field of rational. Then the cyclotomic polynomial is denoted by  $\Phi_p$  and is given by

$$\Phi_p(x) = 1 + x + x^2 + \dots + x^{p-1}$$

(a)

$$\text{Let } f(x) = (x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n)$$

Now consider the derivative of above function

Choose  $(x - \alpha_1)$  as the first function and the rest of the expression as second then

$$\frac{d(f(x))}{dx} = (x - \alpha_1) [(x - \alpha_2) \dots (x - \alpha_n)]' + ((x - \alpha_2) \dots (x - \alpha_n))$$

Now at  $x = \alpha_1$

$$f'(\alpha_1) = (\alpha_1 - \alpha_2) \dots (\alpha_1 - \alpha_n)$$

In order to evaluate  $f'(\alpha_2)$ , choose  $(x - \alpha_2)$  as the first function and the rest of the expression as the second, then

$$\frac{d(f(x))}{dx} = (x - \alpha_2) [(x - \alpha_1) \dots (x - \alpha_n)]' + ((x - \alpha_1) \dots (x - \alpha_n))$$

Now at  $x = \alpha_2$

$$f'(\alpha_2) = (\alpha_2 - \alpha_1) \dots (\alpha_2 - \alpha_n)$$

In general  $f'(\alpha_i)$ ,  $1 \leq i \leq n$  is given by

$$f'(\alpha_i) = (-1)^i (\alpha_i - \alpha_1) \dots (\alpha_i - \alpha_{i-1}) (\alpha_{i+1} - \alpha_i) \dots (\alpha_n - \alpha_i), \text{ where } i \neq j$$

Then the expression  $f'(\alpha_1) f'(\alpha_2) \dots f'(\alpha_n)$  is given by

$$\prod_{i=1}^n f'(\alpha_i) = (-1)^{1+2+\dots+n} \left( \prod_{i < j} (\alpha_i - \alpha_j)^2 \right)$$

Now since  $1 + 2 + 3 + \dots + n = \frac{n(n-1)}{2}$

Thus,

$$\begin{aligned}\prod_{k < j} (\alpha_k - \alpha_j)^2 &= \left((-1)^{1+2+\dots+n}\right) \prod_{i=1}^n f'(\alpha_i) \\ &= (-1)^{\frac{n(n-1)}{2}} \left(\prod_{i=1}^n f'(\alpha_i)\right)\end{aligned}$$

So the sign depends upon the degree of the function  $f(x)$ .

**Therefore, the result stated in the question has been proved.**

**(b)**

Let  $f(x) = x^p - 1$

Then the discriminant of the function is given as shown in part (a)

So,

$$f'(x) = px^{p-1}$$

Now there are  $p$  roots of the polynomial  $x^p - 1$  given by the set

$$A = \{1, \zeta, \zeta^2, \zeta^3, \zeta^4, \dots, \zeta^{p-1}\}, \text{ where } \zeta = e^{2\pi i/p}$$

Then,

$$\Delta = f'(1)f'(\zeta)f'(\zeta^2)f'(\zeta^3)f'(\zeta^4)f'(\zeta^5)f'(\zeta^6)f'(\zeta^7)$$

So, for every  $1 \leq i \leq p-1$

$$f'(\zeta^i) = p(\zeta^i)^{p-1}$$

Thus,

$$\begin{aligned}\Delta &= \prod_{i=1}^{p-1} \left(p(\zeta^i)^{p-1}\right) \\ &= p^p \left(\zeta^{(p-1)}\right)^{\sum_{i=1}^{p-1} i} \\ &= p^p \left(\zeta^{(p-1)}\right)^{(p-1)(p-2)/2} \\ &= p^p \left(\zeta\right)^{(p-1)^2(p-2)/2}\end{aligned}$$

Now let  $\zeta_p$  be a primitive  $p^{\text{th}}$  root of unity

Then,

$$\begin{aligned}x^p - 1 &= (x-1)(1+x+x^2+\dots+x^{p-1}) \\ &= (x-1)\Phi_p(x)\end{aligned}$$

Differentiate both sides with respect to  $x$ , then

$$px^{p-1} = \Phi_p(x) + (x-1)\Phi_p'(x)$$

Then for each  $1 \leq i \leq p-1$

$$p(\zeta_p^i) = (\zeta_p^i - 1)\Phi_p'(\zeta_p^i)$$

Thus,

$$\begin{aligned}\prod_{i=1}^{p-1} \Phi_p(\zeta_p^i) &= \prod_{i=1}^{p-1} \left( p(\zeta_p^i)^{p-1} \right) / (\zeta_p^i - 1) \\ &= \frac{p^{p-1}}{\prod_{i=1}^{p-1} (\zeta_p^i - 1)} \\ &= p^{p-2}\end{aligned}$$

From the above derivative formula

$$\Delta(\Phi_p(x)) = (-1)^{\frac{p(p-1)}{2}} (p^{p-2})$$

Since the Galois group of  $\mathbb{Q}(\zeta_p)$  over the field of rational number is a cyclic group of order  $p-1$  and say is denoted by  $G$ .

Thus, there exists a unique subgroup say  $H$  of the Galois group of  $\mathbb{Q}(\zeta_p)$  over the field of rational number such that

$$[G:H] = 2$$

Then there exists a unique subfield of  $\mathbb{Q}(\zeta_p)$  which is a quadratic extension of  $\mathbb{Q}$

And  $\sqrt{\Delta(\Phi_p(x))} \in \mathbb{Q}(\zeta_p)/\mathbb{Q}$ , so it generates that unique subfield of  $\mathbb{Q}(\zeta_p)$  which is mentioned above.

Now when  $p \equiv 1 \pmod{4}$ , this implies that  $p = 4k + 1$  where  $k \in \mathbb{Z}$

$$\begin{aligned}\sqrt{\Delta(\Phi_p(x))} &= \sqrt{(-1)^{\frac{p(p-1)}{2}} (p^{p-2})} \\ &= \sqrt{(-1)^{\frac{(4k+1)(4k)}{2}} (p^{p-2})} \\ &= \sqrt{(-1)^{2k(4k+1)} (p^{p-2})} \\ &= \sqrt{p}\end{aligned}$$

Similarly, when  $p \equiv 3 \pmod{4}$ , then  $\sqrt{\Delta(\Phi_p(x))} = \sqrt{-p}$

**Therefore, the result stated in the question has been proved.**

10. a



An Eigen vector is defined as the vector which when gets operated on by a given operator gives a scalar multiple of itself.

[Comment](#)

## Step 2 of 3 ^

Let  $F$  be a subfield of  $\mathbb{C}$  that contains  $p^{\text{th}}$  root of unity such that;

$$\zeta = e^{2\pi i/p}$$

Here,  $p$  is a prime

Let  $K$  be a splitting field of  $g$  over  $F$

Now, the Galois group  $G = G(K/F)$  contains an element  $\sigma$  different from the identity.

Let  $\beta$  of  $g$  is not in  $F$

So, define a mapping;

$$\sigma: G \rightarrow G(K/F)$$

Such that;

$$\sigma(\beta) = \zeta^{2v} \beta$$

For some  $v$  with  $0 < v < p$

In general;

$$\sigma^k(\beta) = \zeta^{kv} \beta$$

Now, choose any  $\sigma$  for any cyclic Galois  $G$  then;

$$\sigma^p = 1$$

So, any Eigen Value of  $\sigma$  must satisfy the relation;

$$\lambda^p = 1$$

This means that  $\lambda$  is a power of  $\zeta$

Further, let  $\beta$  be an Eigen vector of  $\sigma$  with  $b = \beta^p$ , then;

$$\sigma(\beta) = \lambda(\beta)$$

And;

$$\begin{aligned} \sigma(b) &= (\lambda\beta)^p \\ &= b \end{aligned}$$

Since,  $\sigma$  is generated by  $G$  and  $b$  is fixed, so;

$$F(\beta) = K$$

Now, consider the roots of  $f$  that yield to an Eigen vector for the operator  $\sigma$  and consider the permutation of roots  $\alpha_1, \dots, \alpha_p$  and let  $\lambda$  be an Eigen value of  $\sigma$  such that;

$$\beta = \sigma_1 + \lambda\sigma_2 + \dots + \lambda^{p-1}\sigma_p$$

Now, replace;

$$\lambda = \zeta^i$$

$$\beta = \gamma_i$$

This implies that;

$$\beta = \sigma_1 + \zeta^i \sigma_2 + \dots + \zeta^{(p-1)i} \sigma_p$$

Here, if  $\beta = 0$  then it will be an Eigen vector with Eigen value  $\lambda^{-1}$ .

**Therefore, at least one of the elements  $\gamma_i = \sigma_1 + \zeta^i \sigma_2 + \dots + \zeta^{(p-1)i} \sigma_p$  is not zero**

# Section 11

## 1. a

Discriminant is defined as the function of the coefficients of a polynomial equation whose values tells about the roots of the polynomial.

To prove: that if the discriminant of an irreducible cubic polynomial in  $F[x]$  is not a square in  $F$ , then the roots cannot be obtained by adjoining a cube root to  $F$

For the proof consider  $a, b, c$  to be independent indeterminates over a field  $K$

Let  $z$  be a zero of the cubic polynomial;

$$x^3 + ax^2 + bx + c$$

Let this polynomial be in some algebraic closure of  $K = K(a, b, c)$

Claim: That  $f(x) = x^3 + ax^2 + bx + c$  is irreducible in  $K(a, b, c)[x]$

Since, a polynomial with coefficients  $x$  in the ring;

$$K(a, b)[c][x] \approx K(a, b)[x][c]$$

But, it is known that;  $[K(z) : K]$  is equal to the degree of the minimal polynomial of  $z$  over  $K$

Since,  $f$  is irreducible it is the minimal polynomial of  $z$  over  $K$ , so;

$$[K(z) : K] = 3$$

Then its determinant will be;

$$\Delta = (z - u)^2 (u - v)^2 (v - z)^2$$

And, this determinant lies in the splitting field and is a square in the splitting field. But if the determinant is not a square in the field  $K$ , then the splitting field consists of the quadratic field

$$K(\sqrt{\Delta}) \text{ and which is of degree 2 over } K$$

Now, use the fact that,  $a, b, c$  are indeterminates and if the characteristic of  $K$  is not 2, then, map;

$$a \rightarrow 0$$

$$c \rightarrow 0$$

So, that  $f(x)$  becomes  $x^3 + bx$

The zeroes of this polynomial will become 0 and  $\pm\sqrt{b}$ , so, the discriminant is;

$$\begin{aligned} \Delta &= (0 - \sqrt{b})^2 \times (0 + \sqrt{b})^2 \times (-\sqrt{b} - \sqrt{b})^2 \\ &= b \times b \times 4b \\ &= 4b^3 \\ &= (2b^2)b \end{aligned}$$

The indeterminate  $b$  is not a square.

That is since this map is not a square; the discriminant is not a square in  $K(a, b, c)$

But also, in characteristic 2, separable quadratic extensions are not all obtained in square roots rather adjointed by a cube root.

**Therefore, if the discriminant of an irreducible cubic polynomial in  $F[x]$  is not a square in  $F$ , then the roots cannot be obtained by adjoining a cube root to  $F$**

## 2. a

a.

Let  $F \subseteq G$  be a splitting for  $f$

Suppose  $\beta \in G$  a root of  $f$

Then, this implies that;

$$\beta^p = b$$

That is;

$$\begin{aligned} f &= x^p - b \\ &= x^p - \beta^p \\ &= (x - \beta)^2 \end{aligned}$$

Let  $f = g_1 \dots g_k$  be a factorization of  $f$  in  $F[x]$  into monic irreducible polynomial and has  $\beta$  as its root.

Thus, each;

$$g_i = n_{F,\beta}$$

This implies the degree of  $n_{F,\beta}$  divides  $p$

Thus,  $f$  is irreducible over  $F$

This implies that either;

$$f = n_{F,\beta}$$

Or;

$$n_{F,\beta} = x - \beta$$

**Therefore, for every  $\beta \in F$ ,  $f$  splits over  $F$**

Further, let  $\beta$  be the root of  $f$  in  $G$  such that;

$$G = G[\beta, \epsilon]$$

As, each root in  $f$  is of the form  $\epsilon^k \beta$

This says that  $f$  splits over  $G$

Clearly, it is known that;

$$[E : \mathbb{Q}[\epsilon]] = p$$

Also,  $G$  is Galois over  $G \cap F$ . So, by definition of towers;

$$\begin{aligned} [F[\beta] : F] &= |(F, G) : F| \\ &= |G : G \cap F| \end{aligned}$$

This expression divides;

$$[G : \mathbb{Q}[\epsilon]] = p$$

Then,  $G \subseteq F$

This implies that  $f$  splits over  $F$

**Thus,  $f$  is irreducible**

b.

Consider the characteristic of  $f$  to be  $p$

Suppose  $f$  is irreducible defined as;

$$f = a^p$$

Where,  $\beta$  is a root of  $a$

Then, in  $K(\beta)[x]$ :

$$f = (x - \beta)^p$$

This means;

$$a = (x - \beta)^n$$

Also, all other factors of  $f$  are of the same form.

So, the degree of the root will be 1 or  $p$  for all factors of  $f$

Since, the characteristic of  $F$  is  $p$

And, the degree of roots in terms of polynomial is 1 or  $p$

Also, it is given that  $x^p - a$  is irreducible  $F[x]$

**Therefore, it has a root in  $F$**

### 3. a

Galois extensions are defined as those extension fields which are both normal and separable.

Consider  $F$  be a subfield of  $\mathbb{C}$  that contains  $i$ , and let  $K$  be a Galois extension of  $F$

Let  $\alpha$  be a generator for the cyclic group  $C_4$ , it will be regarded as an  $F$ -linear transformations of  $K$

Let  $F[\alpha]$  be the ring of  $F$ -linear transformations of  $K$  generated by  $\beta$  whose elements have the form of  $c_0I + c_1\beta + \dots + c_n\beta^n$

Now, take the homomorphism mapping;

$$\varphi: F[x] \rightarrow F[\beta]$$

Such that;

$$\varphi(f(x)) = f(\beta)$$

Where,  $\varphi(1) = I$

Here, the kernel of  $\varphi$  contains  $y^4 - 1$  and the characteristic polynomial  $\det(\beta - yI)$

Hence, the roots of  $f(y)$  are roots of  $y^4 - 1$  and have multiplicity 1 since, the roots  $y^4 - 1$  have multiplicity 1.

Now, since  $p(y)$  divides  $\det(\alpha - yI)$ , the roots of  $f(y)$  are  $\beta$ -Eigen values

Thus, it can be seen that the roots of  $f(x)$  are fourth roots of  $I$  of multiplicity 1 and are Eigen values of  $\beta$

If neither  $i$  nor  $-i$  is a root of  $f(y)$  then  $f(y)$  is one of the following;

$$y-1, y+1$$

That is;

$$(y-1)(y+1) = y^2 - 1$$

This cannot be possible since,  $\beta$  has order 4 not 1 or 2.

Now, suppose then that  $i$  is a root of  $p(y)$  and so  $\beta$ -Eigen value

Then there is a  $\beta$ -Eigen vector  $\alpha \in K$  of value  $i$

The  $G$ -orbit of  $\alpha \in G$  is;

$$\{\alpha, \beta\alpha, \beta^2\alpha, \beta^3\alpha\} = \{\alpha, i\alpha, -\alpha, -i\alpha\}$$

And, so the irreducible polynomial for  $\alpha \in F[x]$  has degree 4

Hence;

$$K = F[\alpha]$$

Also;

$$\begin{aligned}\beta(\alpha^4) &= (\delta\alpha)^4 \\ &= \alpha^4\end{aligned}$$

Therefore,  $\boxed{\alpha^4 \in F}$

**Hence, the statement is true.**

4. a

Consider the provided statement to carry out computation to arrive at Cardano's formula.

As it is known that, Cardano's formula is used to solve the cubic equation.

[Comment](#)

Step 2 of 3 ^

Let is assumed that the cubic equation is  $ax^3 + bx^2 + cx + d = 0, a \neq 0$  where  $a, b, c, d \in \mathbb{Z}$ .

Now divide the standard cubic equation by  $a$  then,

$$x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} = 0$$

Now substitute the value of  $x$  such that  $x = y - \frac{b}{3a}$  then the equation will become,

$$\begin{aligned}x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} &= 0 \\ \left(y - \frac{b}{3a}\right)^3 + \frac{b}{a}\left(y - \frac{b}{3a}\right)^2 + \frac{c}{a}\left(y - \frac{b}{3a}\right) + \frac{d}{a} &= 0 \\ \left(y^3 - \frac{b^3}{27a^3} - 3y^2\left(\frac{b}{3a}\right) + 3y\left(\frac{b^2}{9a^2}\right) + \right. \\ \left. \frac{b}{a}\left(y^2 - 2y\frac{b}{3a} + \frac{b^2}{9a^2}\right) + \left(\frac{c}{a}\right)y - \frac{cb}{3a^2} + \frac{d}{a}\right) &= 0 \\ y^3 + \frac{2b^3}{27a^3} - \frac{b^2y}{3a^2} - \frac{bc}{3a^2} + \frac{cy}{a} + \frac{d}{a} &= 0\end{aligned}$$

This equation can be also written in the form of  $y^3 + uy + v = 0$  where,

$$u = -\frac{b^2}{3a^2} + \frac{c}{a}, v = \frac{d}{a} + \frac{2b^3}{27a^3} - \frac{bc}{3a^2}$$

Therefore the discriminant of the cubic equation is,

$$\Delta = \left(\frac{u}{3}\right)^3 + \left(\frac{v}{2}\right)^2$$

From the Cardano's formulas, the roots of the cubic equation are shown as below;

$$\begin{aligned} y_1 &= \alpha + \beta \\ y_2 &= -\left(\frac{\alpha + \beta}{2}\right) + i\sqrt{3}\left(\frac{\alpha - \beta}{2}\right) \\ y_3 &= -\left(\frac{\alpha + \beta}{2}\right) - i\sqrt{3}\left(\frac{\alpha - \beta}{2}\right) \end{aligned}$$

Where  $\alpha = \left(\sqrt{-\left(\frac{v}{2}\right) + \sqrt{\Delta}}\right)^{\frac{1}{3}}, \beta = \left(\sqrt{-\left(\frac{v}{2}\right) - \sqrt{\Delta}}\right)^{\frac{1}{3}}$  and for each  $\alpha, \beta$  is taken  $\alpha\beta = -\frac{p}{3}$ .

There is some condition from which it is obtained that which types of roots has been obtained.

- (i) If  $\Delta < 0$  then the equation has three real roots.
- (ii) If  $\Delta > 0$  then the equation has one real root and two roots are complex conjugates.
- (iii) If  $\Delta = 0$  then the equation has two real roots and if  $u = v = 0$  then it has one real root.

5. a

(a)

Consider the provided statement to express the roots of polynomials by using Cardano's formula.

As it is known that for the polynomial function  $f(x) = x^3 + 3px + 2q$  the result of this in computation in Cardano's formula is as below:

$$u = \sqrt[3]{-q + \sqrt{q^2 + p^3}} + \sqrt[3]{-q - \sqrt{q^2 + p^3}}$$

Therefore for the given polynomial function  $f(x) = x^3 + 3x$  is  $q = 0, p = 1$  then,

$$\begin{aligned} u &= \sqrt[3]{0 + \sqrt{0^2 + (1)^3}} + \sqrt[3]{0 - \sqrt{0^2 + (1)^3}} \\ u &= \sqrt[3]{0 + \sqrt{1}} + \sqrt[3]{0 - \sqrt{1}} \end{aligned}$$

Hence, the root of the polynomial is expressed as  $\boxed{\sqrt[3]{0 + \sqrt{1}} + \sqrt[3]{0 - \sqrt{1}}}$ .

[Comment](#)

Step 2 of 5 ^

For the polynomial function,  $f(x) = x^3 + 2$  the value of  $p, q$  is as  $p = 0, q = 1$

$$\begin{aligned} u &= \sqrt[3]{-1 + \sqrt{(1)^2 + (0)^3}} + \sqrt[3]{-1 - \sqrt{(1)^2 + (0)^3}} \\ u &= \sqrt[3]{-1 + \sqrt{1}} + \sqrt[3]{-1 - \sqrt{1}} \end{aligned}$$

Hence, the root of the polynomial is expressed as  $\boxed{\sqrt[3]{-1 + \sqrt{1}} + \sqrt[3]{-1 - \sqrt{1}}}$ .

For the polynomial function,  $f(x) = x^3 - 3x + 2$  the value of  $p, q$  is as  $p = -1, q = 1$

$$\begin{aligned} u &= \sqrt[3]{-1 + \sqrt{(1)^2 + (-1)^3}} + \sqrt[3]{-1 - \sqrt{(1)^2 + (-1)^3}} \\ &= \sqrt[3]{-1 + \sqrt{1-1}} + \sqrt[3]{-1 - \sqrt{1-1}} \\ &= \sqrt[3]{-1 + \sqrt{0}} + \sqrt[3]{-1 - \sqrt{0}} \\ &= \sqrt[3]{-1} + \sqrt[3]{-1} \end{aligned}$$

Hence, the root of the polynomial is expressed as  $\boxed{\sqrt[3]{-1} + \sqrt[3]{-1}}$ .



(b)

To find the correct choices of roots in Cardano's formula of the given polynomial function.

For the polynomial function  $f(x) = x^3 + 3x$ , the solutions are  $0, \pm i\sqrt{3}$ . To find the root 0, first of

all choose  $\sqrt[3]{1}$  have same sign to be  $\pm 1$  in both terms and then choose  $\sqrt[3]{\pm 1} = \pm e^{\frac{2\pi ik}{3}}$  and

$$\sqrt[3]{\mp 1} = \mp e^{\frac{2\pi ik}{3}}, k \in \{1, 2\}.$$

For the root  $i\sqrt{3}$  choose  $\sqrt[3]{1}$  to have different signs such as  $\pm 1, \mp 1$  so that,

$$\begin{aligned} u &= \sqrt[3]{0 + \sqrt{1}} + \sqrt[3]{0 - \sqrt{1}} \\ &= 2\sqrt[3]{\pm 1} \\ &= \pm 2\sqrt[3]{1} \end{aligned}$$

Then choose  $\sqrt[3]{1} = e^{\frac{2\pi ik}{3}}, k = 1$ , if choose  $+$  in place of  $\pm$  and  $k = 2$ . If  $-$  is choose in place of  $\pm$  for the root  $-i\sqrt{3}$  would be switch our choice for  $k$ .

For the polynomial function  $f(x) = x^3 + 2$  then choose the same square root that gives

$$\sqrt[3]{-1 + \sqrt{1}} + \sqrt[3]{-1 - \sqrt{1}} = \sqrt[3]{-2} \text{ and each choice of cube root gives a solution.}$$

For the polynomial function  $f(x) = x^3 - 3x + 2$ , it can be factorized as below

$$x^3 - 3x + 2 = (x - 1)^2 (x + 2)$$

Therefore, from the above calculation it is obtained that  $1, -2$  are the solutions. The root  $-2$  is obtained by choosing  $-1$  for both.  $-1$  is obtained by choosing a primitive cube root of unity  $\alpha$  for one root and  $\bar{\alpha}$  for other root.

## Section 12

1. a

Any Galois group containing a subfield is said to be solvable if it has an abelian subfield of the same degree

[Comment](#)

Step 2 of 3 ^

Consider  $K$  to be a subfield of degree 10

Then this subfield must contain a subfield of degree 5

Let such a subgroup be of the form  $Q(\alpha\alpha')$

This subfield is unique as each order 2 subgroup lies in a unique 2-sylow group

Now, by the conjugacy of 2-sylow and the uniqueness of the Sylow, it remains to check that the centralizer of the order 2 element  $\sigma^2$  is  $\langle \sigma \rangle$

Now, the degree 10 extension has  $Q(\alpha\alpha')$  as its unique subfield of degree 5 over  $Q$  and has

$Q(\sqrt{5})$  as its unique subfield.

Now, Galois group  $G$  is said to be solvable if there exists a chain of subgroups;

$$G_0 \supseteq G_1 \supseteq \dots \supseteq G_n$$

Such that;

$$G_{n+1} \leq G_n$$

And;

$$G_n/G_{n+1}$$

This should be an abelian group

Since, every Abelian group is solvable. So, it is enough to show that subfield of degree 10 is abelian.

For this it is enough to show that subfield of degree 5 is abelian or not.

Take any subfield of the form of;

$$x^5 - 2x + 2$$

Now, this expression is not abelian.

So, clearly it can be said that subfield of degree 5 is non-abelian.

**Therefore, Galois group 10 is not solvable.**

## 2. a

Consider the provided statement to determine the transitive subgroup of  $S_5$ .

As it is known that, a subgroup  $G$  of  $S_5$  is transitive if and only if 5 divides the order of  $G$  and if and only if  $G$  contains a 5-cycle.

[Comment](#)

### Step 2 of 3 ^

Any group which is irreducible polynomial of degree 5 is isomorphic to a transitive subgroup of  $S_5$  and their order is a multiple of 5. Therefore only divisors of  $|S_5| = 120$  that are multiples of 5 are 5, 10, 15, 20, 30, 40, 60 and 120.

By Cauchy's theorem,  $G$  contains an element of order 5. The only elements  $S_5$  of order 5 are the 5-cycles. However, if  $\sigma$  is a 5-cycle, then for each  $i, j \in \{1, 2, 3, 4, 5\}$ . There is a power of  $\sigma$  which send  $i$  to  $j$ . Therefore  $G$  is transitive.

The transitive groups of degree 5 is shown As below;

Degree	Order	Description	Generators
5	5	$C_5$	$(12345)$
5	10	$ASL_4(5)$	$(12345), (25)(34)$
5	20	$ASL_4(5)$	$(12345), (2354)$

Therefore, the transitive subgroups of  $S_5$  are  $\{S_5, A_5\} \cup \{\langle \sigma \rangle, U_\sigma, V_\sigma \mid \sigma \text{ a 5-cycle}\}$  in which  $U_\sigma$  are conjugated and all the  $V_\sigma$  are conjugated.

## 3. a

Consider the provided statement to prove that the group is either  $S_5$  or  $A_5$ .

It is assumed that  $K$  is a splitting field of  $f$ . If  $G = S_5$  then the discriminant of  $f$  is not a square in  $F$ . If  $F$  is replaced by  $F(\delta)$  where  $\delta$  is a square root of the discriminant in  $K$ .

The Galois group of  $G\left(\frac{K}{F(\delta)}\right)$  is in  $A_5$ .

[Comment](#)

#### Step 2 of 2 ^

It is assumed that the Galois group of  $f$  is  $A_5$  but that some root  $\alpha$  of  $f$  is expressible by radicals over  $F$ . It can be said that  $\alpha \in F_r$  where  $F = F_0 \subset \dots \subset F_r$  and every extension in the chain is Galois with cyclic Galois group.

As the Galois group of  $f$  over  $F$  is a simple group that is  $A_5$ . From the previous theorem, it can be seen that the Galois group of  $f$  over  $F_i$  is  $A_5, \forall i$ .

As since  $\alpha \in F_r$  then it is seen that  $f$  is not irreducible over  $F_r$  therefore in particular the Galois group of  $f$  over  $F_r$  is not act transitively on the five roots of  $f$  in the splitting field.

Hence, if an element of order 3 then it is either  $S_5$  or  $A_5$  is **proved**.

## 4. a

Every symmetric polynomial  $g(u_1, \dots, u_n)$  with coefficients in a ring  $R$  can be written in a unique way as a polynomial in the elementary symmetric function  $s_1, s_2, \dots, s_n$ .

Let  $G$  be a group and if following conditions hold

1. A subgroup of  $G$  is isomorphic to  $S_n$
2.  $G$  is isomorphic to a subgroup of  $S_n$

Then,  $G \cong S_n$

(a)

Let  $F$  be a field

Let  $F(u)$  be field of rational functions in  $u_1, u_2, \dots, u_n$ .

Then,  $F(u) = F(u_1, u_2, \dots, u_n)$

Let  $s_1, s_2, \dots, s_n$  be elementary symmetric functions in variables  $u_1, u_2, \dots, u_n$

Consider the polynomial  $g(x)$  in  $F(s_1, s_2, \dots, s_n)[x]$  given by

$$g(x) = x^n - (u_1 + u_2 + \dots + u_n)x^{n-1} + (u_1u_2 + u_1u_3 + \dots + u_{n-1}u_n)x^{n-2} + \dots + (-1)^n (u_1u_2u_3 \dots u_n)$$

Then,  $g(x)$  can be re-written as

$$g(x) = x^n - s_1x^{n-1} + s_2x^{n-2} - \dots + (-1)^n s_n$$

$$g(x) = (x - u_1)(x - u_2)(x - u_3) \dots (x - u_n)$$

Clearly, the above polynomial splits completely in the field  $F(u_1, u_2, \dots, u_n)$

So,  $F(u_1, u_2, \dots, u_n)$  behaves as the splitting field of the polynomial  $g(x)$  over the field

$$F(s_1, s_2, \dots, s_n).$$

Since the polynomial  $g(x)$  has no repeated roots, so  $g(x)$  is separable.

Hence,  $F(u_1, u_2, \dots, u_n)$  behaves as the splitting field of a separable polynomial over

$$F(s_1, s_2, \dots, s_n),$$

This implies that  $F(u_1, u_2, \dots, u_n)$  is the Galois extension of the field  $F(s_1, s_2, \dots, s_n)$

Consider the group  $S_n$

Let  $S_n$  acts on the field of rational function by permuting the variables

Let  $\sigma \in S_n$  and assume that  $\sigma(u_i) = u_{\sigma(i)}$

Since,  $F(s_1, s_2, \dots, s_n)$  is field of symmetric polynomial, so permuting the variables does not change elements of  $F(s_1, s_2, \dots, s_n)$

This implies that  $\sigma$  does not change  $F$

Hence it can be concluded that  $\sigma \in S_n$  fixes  $F(s_1, s_2, \dots, s_n)$

Since  $\sigma \in S_n$  was arbitrary, so every permutation fixes  $F(s_1, s_2, \dots, s_n)$ .

This implies that Galois group of  $F(u_1, u_2, \dots, u_n)/F(s_1, s_2, \dots, s_n)$  denoted by  $G$  has a subgroup say  $H$  such that

$$H \cong S_n$$

Conversely,

Let  $a \in G$  be arbitrary

Then,  $a$  sends  $u_i, 1 \leq i \leq n$  to a root of  $g(x)$

Since roots of  $g(x)$  are  $u_1, u_2, \dots, u_n$ , so assume that  $a$  sends  $u_i$  to  $u_j$

Thus a permutation  $\tau$  can be defined in the following way,

$$\tau(i) = j, 1 \leq i, j \leq n$$

Since,  $a \in G$  was arbitrary

So every element of the group  $G$  can be related to a permutation of  $u_i$ 's

This implies that  $G$  is isomorphic to a subgroup of  $S_n$

Now since  $G$  is isomorphic to a subgroup of  $S_n$  and  $G$  has a subgroup  $H$  such that  $H$  is isomorphic to  $S_n$ .

This implies that  $G \cong S_n$

**(b)**

$$\text{Let } w = u_1u_2 + u_2u_3 + u_3u_4 + u_4u_5 + u_5u_1$$

Consider the polynomial  $g(x)$  as defined in the above part (a)

$$g(x) = x^n - s_1x^{n-1} + s_2x^{n-2} - \dots + (-1)^n s_n$$

Here,

$$s_2 = \sum_{i < j} u_i u_j$$

Then,

$$s_2 = u_1u_2 + u_1u_3 + u_1u_4 + u_1u_5 + u_2u_3 + u_2u_4 + u_2u_5 + u_3u_4 + u_3u_5 + u_4u_5$$

$$\text{Since, } w = u_1u_2 + u_2u_3 + u_3u_4 + u_4u_5 + u_5u_1$$

$$\text{So, } s_2 = w + p, \text{ where } p \text{ denotes } u_1u_3 + u_1u_4 + u_2u_4 + u_3u_5 + u_2u_5$$

Then  $g(x)$  can be re-written in terms of  $w$  and elementary symmetric functions as

$$\begin{aligned} g(x) &= x^n - s_1x^{n-1} + (w+p)x^{n-2} - \dots + (-1)^n s_n \\ &= x^n - s_1x^{n-1} + (w+s_2-w)x^{n-2} - \dots + (-1)^n s_n \\ &= x^n - s_1x^{n-1} + wx^{n-2} + (s_2-w)x^{n-2} - \dots + (-1)^n s_n \end{aligned}$$

So use the result in part (a) to conclude that for  $n=5$  the Galois group for  $F(u)$  over the field  $F(s, w)$ , where  $w = u_1u_2 + u_2u_3 + u_3u_4 + u_4u_5 + u_5u_1$  is isomorphic to  $S_5$ .

Therefore, the Galois group for the field mentioned in the question is  $S_5$ .

(c)

Let  $|G| = n$ ,

Let  $F$  be any field

Let  $F(u_1, \dots, u_n) = K$  be a field in  $n$  variables  $u_1, u_2, \dots, u_n$

Then according to the symmetric function theorem,

Any symmetric polynomial  $h(u_1, \dots, u_n)$  in  $F(u_1, \dots, u_n)$  can be expressed as polynomial in the elementary symmetric functions

So,

$$F(u_1, \dots, u_n)^{S_n} = F(s_1, s_2, \dots, s_n)$$

Thus the field  $F(u_1, \dots, u_n) = K$  is Galois over  $F(s_1, s_2, \dots, s_n)$  as shown in part (a)

Here,  $\text{Aut}_{F(s_1, s_2, \dots, s_n)}(F(u_1, \dots, u_n)) \cong S_n$ .

Now by Cayley's Theorem, every finite group can be embedded into a symmetric group with choice of  $n$

So,  $G$  is isomorphic to a subgroup  $Q$  of  $S_n$ .

Let  $L_Q$  denote the fixed field of  $Q$  in  $K$ .

Then by Fundamental theorem of Galois theory,  $K/L_Q$  is Galois and  $\text{Aut}_{L_Q}(K) = H \cong G$

**Therefore, every finite group is a Galois group for some field and its Galois extension.**

## 5. a

Consider the provided statement to prove that elements of  $K$  can be constructed by ruler and compass.

As it is provided that  $K$  is a Galois extension of  $\mathbb{Q}$  which degree is power of 2. If it is to be shown that any element which has degree 2 can be constructed by ruler and compass. Now it is to be proving that  $(\sqrt{2})^{\frac{1}{3}}$  cannot be constructed with ruler and compass.

[Comment](#)

Step 2 of 2 ^

As it is provided that there is a finite extension  $K$  of  $\mathbb{Q}$  that contains  $(\sqrt{2})^{\frac{1}{3}}$  and it is also provided that  $[K : \mathbb{Q}] = 2^n$ . Then evidently it can be said that  $\mathbb{Q}(\sqrt{2})^{\frac{1}{3}} \subset K$ .

As  $\sqrt[3]{2} \in K$  therefore  $\sqrt[3]{4} = (\sqrt[3]{2})^2 \in K$  is closed under products. Hence, any sums of their rational multiplies also belong to  $K$ .

From the theorem,

$$\begin{aligned} 2^n &= [K : \mathbb{Q}] \\ &= [K : \mathbb{Q}(\sqrt[3]{2})] \cdot [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] \\ &= [K : \mathbb{Q}(\sqrt[3]{2})] \cdot 3 \end{aligned}$$

Therefore, from proposition 16 as 3 divides  $2^n$  is not possible. Hence, elements of  $K$  can be constructed by ruler and compass which degree is power of 2 is **proved**.

## 6. a



**Given:** Given  $f$  is a polynomial and the Galois group of  $f$  is a non-abelian simple group.

**To Prove:** Roots of  $f$  are not solvable.

**Proof:** Let us consider that  $f$  is irreducible. Now recall that if  $\mathbb{F}$  be a subfield of the complex numbers,  $\alpha$  is called solvable over  $\mathbb{F}$  if there is a chain of subfields

$$\mathbb{F} = \mathbb{F}_0 \subset \mathbb{F}_1 \subset \dots \subset \mathbb{F}_r = \mathbb{K}$$

of  $\mathbb{C}$  such that  $\alpha$  is in  $\mathbb{K}$  and for  $j = 1, \dots, r$ ,  $\mathbb{F}_{j+1}$  a Galois extension of  $\mathbb{F}_j$  of prime degree.

This follows that either all roots of  $f$  are solvable or no roots of  $f$  are solvable.

Now we all know that the group  $A_5$  is a simple non-abelian group of order 60 and it contains an element of order 5.

Now consider an arbitrary non-abelian simple group  $G$  instead of  $A_5$ , then we have any sequence of the above form

$$\mathbb{F} = \mathbb{F}_0 \subset \mathbb{F}_1 \subset \dots \subset \mathbb{F}_r = \mathbb{K}$$

terminates in a field  $\mathbb{K}$  over which the Galois group of  $f$  is still  $G$ .

This follows that  $f$  does not split completely over  $\mathbb{K}$ .

Now by the **Splitting Theorem** it yields that  $f$  does not have any root in  $\mathbb{K}$ , therefore no root of  $f$  is solvable.

Now consider that  $f$  is any polynomial (irreducibility condition of  $f$  is dropped now).

Let us assume that  $K$  is the splitting field of  $f$  and let  $\alpha \in \mathbb{C}$  be a primitive element for the extension  $\mathbb{K}/\mathbb{Q}$ .

Let us consider  $g$  be the minimal polynomial for  $\alpha$ . Since  $g$  is irreducible, we can apply the result we derived above for irreducible polynomials that  $g$  has no solvable root and this follows that  $\alpha$  is not solvable.

Now some roots of  $f$  are solvable but at least one root of  $f$  must be not solvable.

Therefore, at least some elements of  $\mathbb{K}$  are not solvable, which implies that the roots of  $f$  cannot all be solvable.

This completes the proof.

**Splitting Theorem:** Let  $\mathbb{K}$  be an extension of a field  $\mathbb{F}$  that is a splitting field of a polynomial  $f(x)$  with coefficients in  $\mathbb{F}$ . If an irreducible polynomial  $g(x)$  with coefficients in  $\mathbb{F}$  has one root in  $\mathbb{K}$ , then it splits completely in  $\mathbb{K}$ .

## Result

3 of 3

Considering an arbitrary non-abelian simple group  $G$  with the sequence of  $\mathbb{F}$  and applying Splitting Theorem we have proved that roots of  $f$  are not solvable.

## 7. a

Consider the provided statement to find a polynomial of degree 7 over  $\mathbb{Q}$  whose Galois group is  $S_7$ .

It is claim that  $S_p$  is generated by a  $p$ -cycle and a transposition for  $p$  prime. Let it is assume that  $(ab \dots cd), (ef)$  be the  $p$ -cycle and transposition. When renaming as  $e=1$  then it is assume that,

$$(ab \dots 1 \dots cd) = (1 \dots cdab), (1f)$$

Therefore,  $(1 \dots cdab \dots)^k = (1f \dots)$  where the range of  $k$  is as  $1 \leq k \leq p-1$  since  $p$  is prime.



Therefore, the generators are  $(1f \dots), (1f)$  by renaming  $f = 2$  and rest of the elements in  $(1f \dots)$  accordingly, the generators are as  $\alpha = (12 \dots p), \beta = (12)$ . Now,

$$\begin{aligned}\alpha^{-1}\beta\alpha &= (1p) \\ \alpha^{-1}(1p)\alpha &= ((p-1)p) \\ \alpha^{-1}((p-1)p)\alpha &= ((p-2)(p-1)), \dots\end{aligned}$$

So, it can generate all permutations of the form  $((k-1)k)$  and these generate  $S_p$ . There is a function  $f(x)$  such that,

$$\begin{aligned}f(x) &= (x^3 - 2)(x^2 - 4)(x^2 - 32) + 2 \\ &= x^7 - 36x^5 - 2x^4 + 128x^3 + 72x^2 - 254\end{aligned}$$

From the Eisenstein criterion, it is irreducible and it has 5 real roots and 2 complex roots. The permutations of its roots that fix the real roots in conjugation that corresponds to a transposition in  $G$ , as  $G$  operates transitively on the roots hence it contains a 7-cycle and the Galois group. Hence, the obtained group  $S_7$  is proved.

## 8. a

Consider the provided statement to prove that the symmetric group  $S_p$  is generated by any  $p$ -cycle together with any transposition.

[Comment](#)

### Step 2 of 3 ^

It is assumed that  $a$  and  $b$  is elements of the symmetric group  $S_p$  where  $a$  has order  $p$  and  $b$  is a transposition then  $\{a, b\}$  generates  $S_p$ .

Now without loss of generality let  $b = (0, 1)$  and as  $a$  has order  $p$  and  $p$  is prime then it is to be proved that  $a$  is  $p$ -cycle. Therefore,  $a^k = (01 \dots), \exists k$

Now re-index the other elements so that  $a^k = (01 \dots p-1)$ . Let  $c = a^k$  then,

$$\begin{aligned}cbc^{-1} &= (01 \dots p-1)(01)(p-1 \dots 01) \\ &= (0)(12)(3) \dots (p-1) \\ &= (12)\end{aligned}$$

Now from induction method,

$$\begin{aligned}c^k bc^{-k} &= c(c^{k-1} bc^{-k+1})c^{-1} \\ &= (k+1, k+2)\end{aligned}$$

Therefore, as  $= (01), (12), \dots, (p-2, p-1)$  are generated by  $\{a, b\}$ .

Let  $(xy)$  be a transposition then,  $(x, x+1)(x+1, x+2) \dots (y-1, y) = (x, y)$  and  $(x, y)$  is also generated by  $\{a, b\}$ . As every permutation can be decomposed into transpositions, therefore it is concluded that  $\{a, b\}$  generates  $S_p$  is **proved**.

## Miscellaneous Problem

### 1. a

Galois extensions are defined as those extension fields which are both normal and separable.

[Comment](#)

Step 2 of 2 ^

Consider  $F_1 \subset F_2$  be field extension.

If  $K_1$  is the splitting field of  $F_1 \in F[x]$  and  $K_2$  is the splitting field of  $F_2 \in F[x]$  and  $K_2$  is the splitting field of  $F_2$

Then since,  $K_1 K_2$  is the smallest subfield of  $K$  generated by the roots of  $F_1$  and  $F_2$

Hence, the splitting field of  $F_1 F_2 \in F[x]$

That is;

$$F_1 \subset F_2 \in F[x]$$

Now, if every irreducible polynomial  $f(x) \in F[x]$  that has a root in  $K_1 \cap K_2$  splits completely in  $G$

Thus,  $g(x)$  splits completely in  $K$  and by the uniqueness of the factorization in  $K[x]$  of  $g$

Therefore,  $g[x]$  splits completely in  $K_1 \cap K_2$

Since, it is given that  $F_1 \subset F_2$

And, clearly  $G(K_1/F_1)$  and  $G(K_2/F_2)$  are Galois extensions

Therefore,  $G(K_1/F_1) \subset G(K_2/F_2)$

2. a

Consider  $F \subset K \subset L$

And, let  $L/F$  and  $K/L$  be Galois extensions

To show: that whether  $K/F$  is necessarily a Galois extension or not

For the proof consider;

$$K = \mathbb{Q}(\sqrt[3]{2})$$

Let  $L$  be the splitting field over  $\mathbb{Q}$  of  $x^3 - 2$

Then  $K/\mathbb{Q}$  will not be a Galois Extensions.

Since, it contains only one root of the irreducible polynomial;

$$x^3 - 2 \in \mathbb{Q}[x]$$

Here,  $K$  is the subfield of  $\mathbb{R}$  but the other roots of  $x^3 - 2$  are not the real numbers.

**Therefore,  $K/F$  will not be necessarily a Galois extension.**

3. a

(a)

Consider the following matrix:

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \dots & u_1^{n-1} \\ 1 & u_2 & u_2^2 & \dots & u_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_n & \dots & \dots & u_n^{n-1} \end{bmatrix}$$

To prove that the determinant of matrix is a constant multiple of the square root of the discriminant  $\delta(u) = \prod_{i < j} (u_i - u_j)$

[Comment](#)

Step 2 of 8 ^

Use the principal of mathematical induction to compute the determinant.

For  $n = 2$ ,

$$\det \begin{pmatrix} 1 & u_1 \\ 1 & u_2 \end{pmatrix} = u_2 - u_1 \\ = \prod_{1 \leq i < j \leq 2} u_j - u_i$$

Suppose that the result is true for all  $n \geq 2$ .

$$\det \begin{pmatrix} 1 & u_1 & u_1^2 & \dots & u_1^{n-1} \\ 1 & u_2 & u_2^2 & \dots & u_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_n & \dots & \dots & u_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (u_j - u_i)$$

For any  $x_1, \dots, x_n \in \mathbb{R}$

[Comment](#)

Step 4 of 8 ^

Now show that this is true for  $n+1$ .


Fix  $u_1, \dots, u_{n+1} \in \mathbb{R}$  and consider  $(n+1) \times (n+1)$  matrix.

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \dots & u_1^{n+1} \\ 1 & u_2 & u_2^2 & \dots & u_2^{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_n & \dots & \dots & u_n^{n+1} \end{bmatrix}$$

By subtracting  $x_1$  times the  $n^{th}$  column of this matrix to the  $(n+1)^{th}$  row, to obtain the matrix,

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \dots & u_1^{n-1} & 0 \\ 1 & u_2 & u_2^2 & \dots & u_2^{n-1} & u_2^n - u_1 u_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & u_{n+1} & u_{n+1}^2 & \dots & u_{n+1}^{n-1} & u_{n+1}^n - x_1 u_{n+1}^{n-1} \end{bmatrix}$$

[Comment](#)

Step 6 of 8 

By subtracting  $x_1$  times the  $(n-1)^{th}$  column of this matrix to the  $n^{th}$  column, to obtain the matrix,

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \dots & u_1^{n-2} & 0 & 0 \\ 1 & u_2 & u_2^2 & \dots & u_2^{n-2} & u_2^{n-1} - u_1 u_2^{n-2} & u_2^n - u_1 u_2^{n-1} \\ 1 & u_3 & u_3^2 & \dots & u_3^{n-2} & u_3^{n-1} - u_1 u_3^{n-2} & u_3^n - u_1 u_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & u_{n+1} & u_{n+1}^2 & \dots & u_{n+1}^{n-2} & u_{n+1}^{n-1} - u_1 u_{n+1}^{n-2} & u_{n+1}^n - u_1 u_{n+1}^{n-1} \end{bmatrix}$$

By continuing this procedure, to obtain,

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & u_2 - u_1 & u_2^2 - u_1 u_2 & \dots & u_2^{n-1} - u_1 u_2^{n-2} & u_2^n - u_1 u_2^{n-1} \\ 1 & u_3 - u_1 & u_3^2 - u_1 u_3 & \dots & u_3^{n-1} - u_1 u_3^{n-2} & u_3^n - u_1 u_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & u_{n+1} - u_1 & u_{n+1}^2 - u_1 u_{n+1} & \dots & u_{n+1}^{n-1} - u_1 u_{n+1}^{n-2} & u_{n+1}^n - u_1 u_{n+1}^{n-1} \end{bmatrix}$$

Therefore, the determinant is,

$$\det \begin{bmatrix} u_2 - u_1 & u_2^2 - u_1 u_2 & \dots & u_2^{n-1} - u_1 u_2^{n-2} & u_2^n - u_1 u_2^{n-1} \\ u_3 - u_1 & u_3^2 - u_1 u_3 & \dots & u_3^{n-1} - u_1 u_3^{n-2} & u_3^n - u_1 u_3^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ u_{n+1} - u_1 & u_{n+1}^2 - u_1 u_{n+1} & \dots & u_{n+1}^{n-1} - u_1 u_{n+1}^{n-2} & u_{n+1}^n - u_1 u_{n+1}^{n-1} \end{bmatrix}$$

Since the determinant is linear in each row. So the determinant of the above matrix is,

$$\prod_{1 \leq j < j' \leq n+1} (u_{j'} - u_j)$$

So, determinant of Vander mode is:

$$\delta(u) = \prod_{i < j} (u_i - u_j)$$

**Hence proved**

(b)

To determine the constant;

Since one term in the determinant is the product of the diagonal elements, which occurs with coefficient  $+1$  and it has the property that the highest possible power  $u_n$  and the next highest possible power is  $u_{n-1}$ , and so on.

In expanding the polynomial product, if prefer then the powers can only be maximized in this way, the term with the largest index.

Since, the form of the sum is always the positive term.

Thus, the coefficient of the diagonal product is constant  $\boxed{+1}$ .

4. a

Consider the provided statement to prove that the field has no automorphism except the identity.

As any automorphism  $\varphi$  is clearly the identity on the integers and rational because,

$$\begin{aligned}\varphi(n) &= \varphi(1 + \dots + 1) \\ &= \varphi(1) + \dots + \varphi(1) \\ &= n\end{aligned}$$

Then,

$$\begin{aligned}q\varphi\left(\frac{p}{q}\right) &= \varphi(p) \\ &= p\end{aligned}$$

Let  $a$  is a number such that  $\varphi(a) = b$  which is not equal to zero. Then  $b < a$  after possibly multiplying by  $-1$ , when subtracting a suitable rational from each then it is find that  $b < 0 < a$ . But then  $a = a^2$  therefore,

$$\begin{aligned}\varphi(a)^2 &= \varphi(a^2) \\ &= \varphi(a) \\ &= b\end{aligned}$$

This provides contradiction as  $b < 0$ .

(b)

From the previous part, any automorphism  $\varphi$  is the identity on  $\mathbb{Q}$ . It is assumed that  $i$  from the definition of complex number such that  $i^2 = -1$ . It is also seen that  $\varphi(i)^2 = -1$ , since only  $\pm i$  square to get  $-1$  in  $\mathbb{C}$  on the  $\mathbb{Q}[i]$ . As  $\varphi$  is either the identity or complex conjugation.

At the end,  $\mathbb{Q}[i]$  is dense in  $\mathbb{C}$ , therefore any automorphism of  $\mathbb{C}$  must also being either the identity or complex conjugation by continuity.

5. a

**Given:**  $K$  is a Galois extension of  $\mathbb{Q}$  whose degree is a power of 2 and  $K \subset \mathbb{R}$ .

**To Prove:** The elements of  $K$  can be constructed by ruler and compass.

**Proof:** Let us first start with a Lemma.

**Lemma:** Let  $\mathbb{Q} = F_0 \subset F_1 \subset \dots \subset F_n = K$  be a chain of subfields of the field  $\mathbb{R}$  of real numbers with the property that for each  $i = 0, 1, \dots, n-1$ ,  $[F_{i+1} : F_i] = 2$ . Then element of  $K$  is constructible.

**Claim:** Suppose  $F$  and  $K$  are fields with

$$\mathbb{Q} \subseteq F \subseteq K \subset \mathbb{R}.$$

Let us consider all elements of  $F$  are constructible, and assume that

$$[K : F] = 2^n, \text{ where } n \text{ is a non-negative integer.}$$

Then, all elements of  $K$  are constructible.

**Proof of the Claim:** We will prove it by using induction on  $n$ .

If  $n = 0$ , then we have  $K = F$  and we are done.

So our first step of induction is over.

Now if  $n = 1$ , then  $[K : F] = 2$ . Therefore by the above lemma our statement is all done.

Now consider  $n \geq 2$ .

Let us assume  $G = \text{Gal}(K/F)$ , a group of order power of 2.

If  $G$  is abelian, then let  $H$  be a subgroup of  $G$  of order 2. Since  $G$  is abelian, every subgroup of  $G$  is normal, hence  $H$  is normal.

Otherwise if  $G$  is not abelian, for our easy let us assume  $H = Z(G)$ .

Clearly  $H$  is a normal subgroup of  $G$  and  $H \neq G$ . Also  $H$  is non-trivial group.

Let us now assume that  $L = K^H$ .

Then by Fundamental Theorem it follows that  $\text{Gal}(K/L) = H$  and  $L/F$  is also a Galois extension with Galois group  $G/H$ .

Now notice that both the groups  $H$  and  $G/H$  have order of the form  $2^m$ , where  $m \in \mathbb{N}$  and satisfying  $m < n$ .

Now applying the induction hypothesis to  $L/F$ , we notice that all elements of  $L$  are constructible, and applying it again to  $K/L$ , we observe that all elements of  $K$  are constructible.

This proves the Claim.

Now notice that in case of our problem,  $F$  is nothing but the field of rational numbers, that is,  $\mathbb{Q}$ .

So by the above claim considering  $F$  as  $\mathbb{Q}$ , we have proved that the elements of  $K$  can be constructed by ruler and compass.

This completes the proof.

## Result

3 of 3

Considering  $K$  is a Galois extension of  $\mathbb{Q}$  whose degree is a power of 2 and  $K \subset \mathbb{R}$  we have proved that elements of  $K$  are constructible.

6. a

(a)

Let  $K = \mathbb{Q}(\zeta)$ , where  $\zeta$  denotes the primitive fifth root of 1

Let  $G$  denotes the group of automorphism on  $K = \mathbb{Q}(\zeta)$

And  $G = \{id, \sigma, \sigma^2, \sigma^3\}$

For,  $K = \mathbb{Q}(\zeta)$ , the group of automorphism is a cyclic group of order 4

Now define  $\sigma: K \rightarrow K$  as

$$\sigma(\zeta) = \zeta^2$$

Or  $\sigma(\zeta) = \zeta^3$

These are the only two possibilities for  $\sigma: K \rightarrow K$

Now for  $\sigma(\zeta) = \zeta^2$

Under the map the following way terms are mapped to each other

$\zeta$	$\zeta^2$
$\zeta^2$	$\zeta^4$
$\zeta^3$	$\zeta$
$\zeta^4$	$\zeta^3$



Now for this map evaluate  $\sigma^2$

$\zeta$	$\zeta^4$
$\zeta^2$	$\zeta^3$
$\zeta^3$	$\zeta^2$
$\zeta^4$	$\zeta$

Now evaluate  $\sigma^3$

$\zeta$	$\zeta^3$
$\zeta^2$	$\zeta$
$\zeta^3$	$\zeta^4$
$\zeta^4$	$\zeta^2$

It can be seen that the another automorphism is also similar to the one studied above

So, restrict the case to only  $\sigma(\zeta) = \zeta^2$

Now consider the pentagon with vertices as  $1, \zeta, \zeta^2, \zeta^3, \zeta^4$

Consider the action of every automorphism on vertices  $1, \zeta, \zeta^2, \zeta^3, \zeta^4$

For  $\sigma$ ,

$$\sigma(1) = 1, \sigma(\zeta) = \zeta^2, \sigma(\zeta^2) = \zeta^4, \sigma(\zeta^3) = \zeta, \sigma(\zeta^4) = \zeta^3$$

Which means the pentagon under the action of  $\sigma$  attains the star shape

Now for  $\sigma^2$

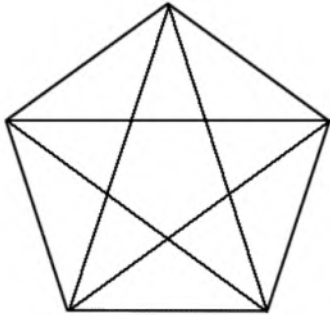
$$\sigma^2(1) = 1, \sigma^2(\zeta) = \zeta^4, \sigma^2(\zeta^2) = \zeta^3, \sigma^2(\zeta^3) = \zeta^2, \sigma^2(\zeta^4) = \zeta$$

Which means the pentagon under the action of  $\sigma^2$  retains the shape but the vertex gets changed or the pentagon gets reversed.

Now for  $\sigma^3$

$$\sigma^3(1) = 1, \sigma^3(\zeta) = \zeta^3, \sigma^3(\zeta^2) = \zeta, \sigma^3(\zeta^3) = \zeta^4, \sigma^3(\zeta^4) = \zeta^2$$

Which means the pentagon under the action of  $\sigma$  attains the star shape



Thus the above shape denotes the  $G$  orbit of regular pentagon with vertices  $1, \zeta, \zeta^2, \zeta^3, \zeta^4$

**Therefore, the result stated in the question has been done**

(b)

Let the length of side of the pentagon be  $\alpha$

Let the five roots of unity are given by

$$1, \frac{1 - \sqrt{5} \pm i\sqrt{10 + 2\sqrt{5}}}{4}, \frac{1 + \sqrt{5} \pm i\sqrt{10 - 2\sqrt{5}}}{4}$$

Now the side of the pentagon can be evaluated by distance formula by considering the points in the complex plane

Thus,

$$\begin{aligned}\alpha &= \sqrt{\left(\left(\frac{1 - \sqrt{5}}{4}\right) - 1\right)^2 + \left(\frac{\sqrt{10 + 2\sqrt{5}}}{4}\right)^2} \\ &= \sqrt{\frac{(3 + \sqrt{5})^2}{16} + \frac{10 + 2\sqrt{5}}{16}} \\ &= \sqrt{\frac{24 + 4\sqrt{5}}{16}}\end{aligned}$$

This implies that  $\alpha^2$  is given by

$$\begin{aligned}\alpha^2 &= \frac{24 + 4\sqrt{5}}{16} \\ &= \frac{3}{2} + \frac{\sqrt{5}}{4}\end{aligned}$$

Now since,

$$\frac{3}{2} \in \mathbb{Q} \subseteq K$$

And  $\mathbb{Q}(\sqrt{5})$  is the unique quadratic extension of  $\mathbb{Q}$  in cyclotomic field  $K$

So,  $\sqrt{5} \in \mathbb{Q}(\sqrt{5})$

This implies that

$$\frac{\sqrt{5}}{4} \in \mathbb{Q}(\sqrt{5}) \subseteq K$$

Since field  $K$  is closed under addition,

Thus,

$$\frac{3}{2} + \frac{\sqrt{5}}{4} \in K$$

Hence,  $\alpha^2 \in K$

Now since the value of  $\alpha^2$  is given by

$$\alpha^2 = \frac{3}{2} + \frac{\sqrt{5}}{4}$$

So, the value of  $\alpha$  is given by

$$\alpha = \sqrt{\frac{3}{2} + \frac{\sqrt{5}}{4}}$$

Square both the sides to obtain the expression

$$\begin{aligned}\alpha^2 &= \frac{3}{2} + \frac{\sqrt{5}}{4} \\ \alpha^2 - \frac{3}{2} &= \frac{\sqrt{5}}{4}\end{aligned}$$

Square both the sides again

$$\left(\alpha^2 - \frac{3}{2}\right)^2 = \left(\frac{\sqrt{5}}{4}\right)^2$$

$$\alpha^4 + \frac{9}{4} - 3\alpha^2 = \frac{5}{16}$$

$$16\alpha^4 - 48\alpha^2 + 31 = 0$$

So, the irreducible equation of  $\alpha$  over  $\mathbb{Q}$  is given by  $16\alpha^4 - 48\alpha^2 + 31$

**Therefore, the result stated in the question has been proved and the irreducible equation of  $\alpha$  over  $\mathbb{Q}$  is given by  $16\alpha^4 - 48\alpha^2 + 31$ .**

7. a

(a)

To be prove that the square root of the discriminant is a skew-symmetric.

It is to be shown that for any permutation of two elements contributes a negative sign, any permutations which is generated by permutations of two elements and the number of permutation of two elements needed to generate a given permutation is equal to its sign.

Therefore interchange the value of  $u_i$  with  $u_j$  without loss of generality such that  $i < j$ .

So, the term in  $\delta$  involving  $u_i, u_j$  and that have the product,

$$(u_i - u_j) \prod_{k=1}^{i-1} (u_k - u_i) \prod_{k=1}^{j-1} (u_k - u_j) \prod_{k=i+1}^{j-1} (u_k - u_i) \prod_{k=j+1}^n (u_i - u_k) \prod_{k=j+1}^n (u_j - u_k) \prod_{k=j+1}^n (u_j - u_k)$$

Now interchange  $u_i, u_j$  that causes a sign change in each factor. The change of sign in  $u_i - u_j$  is the only one that remains. As in each two adjacent factors in the rest of the product have canceling sign changes. Hence,  $\delta$  also changes sign when the value of  $u_i$  and  $u_j$  are interchanged. Hence,  $\delta$  is a skew-symmetric.

(b)

It is assumed that  $h$  be the  $\frac{1}{2}$  symmetric polynomial. Suppose  $\text{char } F \neq 2$  and if  $h$  is symmetric, then let  $f = h, g = 0$  therefore it is not supposed. The action of  $S = S_n$  on  $h$  has orbit  $\{h, h'\}$  for some other polynomial  $h'$ .

Since from the orbit stabilizer theorem,

$$|S_h| |S_h| = |S|$$

But as it is known that  $S_h = A_n$  therefore  $|S_h| = 2$  and this implies that  $f = \frac{1}{2}(h + h')$  is symmetric. The value of  $h, h'$  will be interchanged then,

$$\begin{aligned} x &= h - f \\ &= \frac{1}{2}(h - h') \end{aligned}$$

The above obtained result is anti-symmetric for the same reason.

Now it is to be shown that, any anti-symmetric polynomial  $x$  is divisible by  $\delta$  that is any binomial  $u_i - u_j$  divides it. If  $\varphi$  is the substitution map then it is assumed that  $u_j = u_i$  then  $\varphi(x) = -\varphi(x)$  this implies that  $\varphi(x) = 0$ .

Therefore, letting  $g = \frac{x}{\delta}$  suffices, since  $g$  cannot be anti-symmetric otherwise  $x$  will be symmetric by previous part. Hence, given statement is **proved**.

8. a

The indices of the given quantity define the number of times the given quantity has been multiplied to get the required result.

[Comment](#)

Step 2 of 5 ^

Consider the variables  $u_0, u_1, u_2, u_3$

And define;

$$p_i = (u_i - u_{i+1})(u_i - u_{i+2})(u_{i+1} - u_{i+2})$$

Since, the modulo is 4, so consider the following;

$$u_0 = 0$$

$$u_1 = 1$$

$$u_2 = 2$$

$$u_3 = 3$$

And further;

$$u_4 = 0$$

$$u_5 = 1$$

$$u_6 = 2$$

a.

To determine:  $\sum_{i=0}^3 \frac{u_i}{p_{i+1}}$

First take  $i = 1$

$$\begin{aligned} p_1 &= (u_1 - u_2)(u_1 - u_3)(u_2 - u_3) \\ &= (1 - 2)(1 - 3)(2 - 3) \\ &= -2 \end{aligned}$$

Now, take  $i = 2$ ;

$$\begin{aligned} p_2 &= (u_2 - u_3)(u_2 - u_4)(u_3 - u_4) \\ &= (2 - 3)(2 - 0)(3 - 0) \\ &= -6 \end{aligned}$$

Now, take  $i = 3$ ;

$$\begin{aligned} p_3 &= (u_3 - u_4)(u_3 - u_5)(u_4 - u_5) \\ &= (3 - 0)(3 - 1)(0 - 1) \\ &= -6 \end{aligned}$$

Now, take  $i = 3$ ;

$$\begin{aligned} p_4 &= (u_4 - u_5)(u_4 - u_6)(u_5 - u_6) \\ &= (0 - 1)(0 - 2)(1 - 2) \\ &= -2 \end{aligned}$$

Thus;

$$\begin{aligned} \sum_{i=0}^3 \frac{u_i}{p_{i+1}} &= \frac{u_0}{p_1} + \frac{u_1}{p_2} + \frac{u_2}{p_3} + \frac{u_3}{p_4} \\ &= \frac{0}{-2} + \frac{1}{-6} + \frac{2}{-6} + \frac{3}{-2} \\ &= 0 - \frac{1}{6} - \frac{2}{6} - \frac{9}{6} \\ &= \frac{0 - 1 - 2 - 9}{6} \\ &= \frac{-12}{6} \\ &= -2 \end{aligned}$$

Therefore,  $\boxed{\sum_{i=0}^3 \frac{u_i}{p_{i+1}} = -2}$

b.

To determine:  $\sum_{i=0}^3 \frac{u_i^3}{p_{i+1}}$

First take  $i = 1$

$$\begin{aligned} p_1 &= (u_1 - u_2)(u_1 - u_3)(u_2 - u_3) \\ &= (1-2)(1-3)(2-3) \\ &= -2 \end{aligned}$$

Now, take  $i = 2$ ;

$$\begin{aligned} p_2 &= (u_2 - u_3)(u_2 - u_4)(u_3 - u_4) \\ &= (2-3)(2-0)(3-0) \\ &= -6 \end{aligned}$$

Now, take  $i = 3$ ;

$$\begin{aligned} p_3 &= (u_3 - u_4)(u_3 - u_5)(u_4 - u_5) \\ &= (3-0)(3-1)(0-1) \\ &= -6 \end{aligned}$$

Now, take  $i = 3$ ;

$$\begin{aligned} p_4 &= (u_4 - u_5)(u_4 - u_6)(u_5 - u_6) \\ &= (0-1)(0-2)(1-2) \\ &= -2 \end{aligned}$$

Thus;

$$\begin{aligned} \sum_{i=0}^3 \frac{u_i^3}{p_{i+1}} &= \frac{u_0^3}{p_1} + \frac{u_1^3}{p_2} + \frac{u_2^3}{p_3} + \frac{u_3^3}{p_4} \\ &= \frac{(0)^3}{-2} + \frac{(1)^3}{-6} + \frac{(2)^3}{-6} + \frac{(3)^3}{-2} \\ &= 0 - \frac{1}{6} - \frac{8}{6} - \frac{81}{6} \\ &= \frac{0-1-8-81}{6} \\ &= \frac{-90}{6} \\ &= -15 \end{aligned}$$

Therefore,  $\boxed{\sum_{i=0}^3 \frac{u_i^3}{p_{i+1}} = -15}$

9. a

Any field is defined as the nonzero commutative division ring or can be said as the ring whose nonzero elements form an abelian group under multiplication.

[Comment](#)

Step 2 of 3 ^

Let  $f \in F$ ,  $F$  is a field and let  $K$  be some subring field

Then by primitive element theorem which states that;

Let  $E/F$  be a separable extension of finite degree. Then  $E = F(x)$  for some  $x \in E$  that is the extension is simple and  $x$  is a primitive element

Then by the above stated theorem, there exists a  $a \in K$  such that;

$$K = f(a)$$

Since,  $[K : F]$  is finite

Now, by again using primitive element theorem then;

$$K = f(a_1, \dots, a_n)$$

Take any  $a \in K$  such that  $a$  is a root of  $f(x) \in F[x]$

Thus,  $a$  becomes a root of polynomial  $f(x)$  in  $F[x]$  of degree  $[K : F]$

Now, use the fact that an Automorphism of  $K$  must send  $a$  to another root of  $f(x) \in K$

Consider the field  $\mathbb{Q}(\sqrt[3]{2})$

$$\text{If } f(t_0) = 0$$

Then, the Automorphism of  $\mathbb{Q}(t)$  sends  $a$  to another root to  $f(x)$ , this is true if;

$$t = \sqrt[3]{2}$$

Now,  $\mathbb{Q}(t)/\mathbb{Q}$  is not Galois as  $\mathbb{Q}(t)$  is not a splitting field for any polynomial in  $\mathbb{Q}$

And, since clearly  $t$  is a root of  $x^3 - 2$  and its splitting field has degree 6 over  $\mathbb{Q}$

Thus;

$$[K : F] = 6$$

**Therefore, the splitting field  $K$  of  $f(x)$  over  $\mathbb{Q}(t)$  has degree 6**

10. a

Consider the provided statement to prove the given condition.

It is given that  $K$  is a finite extension of a field  $F$  and  $f(x)$  be in  $K(x)$ . It is assumed that  $f(x)$  is irreducible when it is working with irreducible factors individually and also supposes that it is monic.

[Comment](#)

Step 2 of 2 ^

Let  $\alpha$  is a root of  $f(x)$  which is algebraic over  $K$ . Therefore by previous exercise, it is algebraic over  $F$ .

So  $h(x) \in F[x]$  is the minimal polynomial for  $\alpha$  over  $F$ . As  $f$  is defined over  $K$  then  $f$  divides  $h$  therefore  $h = fg$  where  $g \in K[x]$ . Hence, given statement is **proved**.

11. a



!!!

12. a

Solvable group in a group of fields is defined as the group that can be constructed from the group of abelian using the extensions.

[Comment](#)

Step 2 of 4 ^

Consider  $f$  to be solvable by radicals over  $F$

To show: that  $G$  is a solvable group

For the proof consider  $f_0, f_1, \dots, f_n$  be the roots of  $f$

Then a root tower will be obtained as given below;

$$F = f_0 \subseteq f_1 \subseteq \dots \subseteq f_n$$

Also, it is known that the subgroup of a solvable group is also the solvable one.

Thus, the subgroup  $H$  is solvable

That is;

$$H \subseteq F_n$$

This will be solvable function

And, now by using the theorem which states that;

Let  $f \in F$  be solvable by radicals over  $F$ . Let  $E$  be the splitting field of the solvable group of  $f$  over  $F$ . Then  $G(E/F)$  is a solvable group

Therefore, from the above theorem it can be obtained that  $G$  is solvable group.

Conversely;

Consider  $G$  to be a solvable group.

To show: the roots of  $f$  are solvable

Let  $K$  be the splitting field of  $f(x)$  over  $F$

Let  $f_1, \dots, f_n$  be the roots of  $f$

Clearly, these roots are distinct

Now, assume that the first two terms are non-real.

That is,  $f_1$  and  $f_2$  are not real and  $f_3, \dots, f_n$  are real

Now, clearly,  $G(K/F)$  is a subgroup of  $S_k$

Claim:  $G(K/F) = S_k$

Now, the conjugation restricted to  $K$  in  $G(K/F)$  as a subgroup of  $S_k$  will be of the form of the permutation  $(1, 2)$  of  $S_k$

Thus,  $G(K/F)$  contains a transposition

Further, take out any integer  $a$  and  $b$  with  $1 \leq ab \leq n$

Since,  $f_a$  and  $f_b$  are both roots of the irreducible polynomial  $f$  in  $F$

This there exists a  $\phi \in G(K/F)$  such that;

$$\phi(f_b) = f_a$$

That is  $\phi$  sends  $b$  to  $a$

Since, there is some  $\phi$  for any choice of  $a$  and  $b$

This implies that;

$$G(K/F) = S_k$$

Thus,  $f$  is not solvable by radicals of  $F$

This is a contradiction

**Thus, if  $G$  is solvable group then, the roots of  $f$  are solvable over  $F$**

13. a

Character is defined as the most commonly used special kind of the function which is extracted from the group to a field.

[Comment](#)

Step 2 of 3 ^

Consider  $G$  be a Galois group.

Let  $K/F$  be a Galois extension of the given Galois group  $G$

Consider  $K$  to be a  $F$ -vector space which shows a representation of  $G$  on  $K$

Let  $\chi$  denote the character of this representation.

Now, define a lattice  $L$  such that  $\chi_1, \dots, \chi_n$  are linearly independent over  $L$  then there exists;

$$c_1, \dots, c_n \in L$$

Such that;

$$\sum_{i=1}^n c_i \chi_i(g) = 0 \forall g \in G$$

Here, each  $c_i = 0$

Suppose on the contrary that;

$$f_1\chi_1 + \dots + f_m\chi_m = 0$$

Let this be a linear dependence within the condition that  $m$  is minimal possible, that is  $m$  roots of unity.

Clearly,

$$m \geq 2$$

Then without loss generality;

$$f_1 \neq 0$$

Now, fix any  $g \in G$  such that;

$$\chi_m(g) \neq \chi_1(g)$$

That is in particular;

$$f_1\chi_1(y) + \dots + f_m\chi_m(y) = 0 \forall y \in G$$

That is;

$$f_1\chi_1(gy) + \dots + f_m\chi_m(gy) = 0 \forall y \in G$$

Since,  $\chi_i$  multiplicative so;

$$f_1\chi_1(g)\chi_1(y) + \dots + f_m\chi_m(g)\chi_m(y) = 0 \forall y \in G$$

Now multiply the equation  $f_1\chi_1(g) + \dots + f_m\chi_m(g) = 0 \forall g \in G$  by  $\chi_m(g)$ ;

$$f_1\chi_1(g)\chi_m(g) + \dots + f_m\chi_m(g)\chi_m(g) = 0$$

Now, subtract the above obtained equation from

$$f_1\chi_1(g)\chi_1(y) + \dots + f_m\chi_m(g)\chi_m(y) = 0 \forall y \in G, \text{ that is;}$$

$$\sum_{i=1}^{m-1} f_i (\chi_i(g) - \chi_m(g)) \chi_i(x) = 0 \forall x \in G$$

Since;

$$f_i (\chi_i(g) - \chi_m(g)) \neq 0$$

Thus, there is linear independence between  $\chi_1, \dots, \chi_{m-1}$

This contradicts the minimality of  $m$

And this shows that the character of the required representation

**Therefore,  $F$  contains enough roots of unity than  $\chi$  is the character of the required representation.**

# Appendix: Background Material

---

1. a

(a) We claim that the closed form for this expression with  $n$  terms is  $n^2$ , i.e.,  $1 + 3 + 5 + \dots + (2n + 1) = (n + 1)^2$ . We will prove this by induction. Let  $n = 1$ . Then  $(n + 1)^2 = 4 = 1 + 3$ , so the claim holds. Now assume the claim holds for some  $k \in \mathbb{N}$ , i.e.,

$$1 + 3 + 5 + \dots + 2k + 1 = (k + 1)^2. \quad (1)$$

For  $k + 1$  then it holds that:

$$\begin{aligned} 1 + 3 + 5 + \dots + (2k + 1) + (2k + 3) &\stackrel{(1)}{=} (k + 1)^2 + 2k + 3 \\ &= k^2 + 2k + 1 + 2k + 3 \\ &= k^2 + 4k + 4 \\ &= (k + 2)^2 \end{aligned}$$

Therefore, the claim follows by mathematical induction.

We claim that  $1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$ . Let us denote this claim by  $P_n$ .

Firstly, for  $n = 1$ , it holds that:  $\frac{n(n+1)(2n+1)}{6} = \frac{1 \cdot 2 \cdot 3}{6} = 1 = 1^2$ , so  $P_1$  is indeed true.

Let us assume that  $P_k$  holds for some  $k \in \mathbb{N}$ . For  $k + 1$ , we have:

$$\begin{aligned} 1^2 + 2^2 + \dots + k^2 + (k + 1)^2 &\stackrel{P_k}{=} \frac{k(k+1)(2k+1)}{6} + (k + 1)^2 \\ &= (k + 1) \left( \frac{k(2k+1)}{6} + (k + 1) \right) \\ &= (k + 1) \frac{2k^2 + k + 6k + 6}{6} \\ &= (k + 1) \frac{(k + 2)(2k + 3)}{6} \\ &= \frac{(k + 1)((k + 1) + 1)(2(k + 1) + 1)}{6} \end{aligned}$$

This is exactly  $P_{k+1}$  so the claim is proven by mathematical induction.

---

## Result

In this exercise, we practice mathematical induction on two classic examples.

2. a

We have to show that  $1^3 + 2^3 + \dots + n^3 = \frac{(n(n+1))^2}{4}$ . Let us denote this claim by  $P_n$ .

Firstly, for  $n = 1$ , it holds that:  $\frac{(n(n+1))^2}{4} = \frac{(1 \cdot 2)^2}{4} = 1 = 1^3$ , so  $P_1$  is indeed true.

Let us assume that  $P_k$  holds for some  $k \in \mathbb{N}$ . For  $k + 1$ , we have:

$$\begin{aligned} 1^3 + 2^3 + \dots + k^3 + (k+1)^3 &\stackrel{P_k}{=} \frac{(k(k+1))^2}{4} + (k+1)^3 \\ &= (k+1)^2 \left( \frac{k^2}{4} + (k+1) \right) \\ &= (k+1)^2 \frac{k^2 + 4k + 4}{4} \\ &= (k+1)^2 \frac{(k+2)^2}{4} \\ &= \frac{((k+1)(k+2))^2}{4} \end{aligned}$$

This is exactly  $P_{k+1}$  so the claim is proven by mathematical induction.

## Result

We prove this in a manner very similar to the previous exercise.

### 3. a

We have to show that  $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{n(n+1)} = \frac{n}{(n+1)}$ . Let us denote this claim by  $P_n$ .

Firstly, for  $n = 1$ , it holds that:  $\frac{n}{(n+1)} = \frac{1}{1 \cdot 2}$ , so  $P_1$  is indeed true.

Let us assume that  $P_k$  holds for some  $k \in \mathbb{N}$ . For  $k + 1$ , we have:

$$\begin{aligned} \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{k(k+1)} + \frac{1}{(k+1)(k+2)} &\stackrel{P_k}{=} \frac{k}{(k+1)} + \frac{1}{(k+1)(k+2)} \\ &= \frac{k(k+2) + 1}{(k+1)(k+2)} \\ &= \frac{k^2 + 2k + 1}{(k+1)(k+2)} \\ &= \frac{(k+1)^2}{(k+1)(k+2)} \\ &= \frac{k+1}{k+2} \end{aligned}$$

This is exactly  $P_{k+1}$  so the claim is proven by mathematical induction.

## Result

This exercise is another standard example of a proof by mathematical induction.

### 4. a

We will prove this claim by induction on  $n = |T|$ . Let us denote the set  $T$  with  $n$  members as  $T_n = \{t_1, t_2, \dots, t_n\}$ .

First, let  $n = 1$ . Assume  $|S| < 1$ . Then  $S$  is the empty set. This is a contradiction with the definition of the surjection, which implies existence of an element in  $S$  mapped to  $t_1$ .

Assume that the claim holds for some  $k \in \mathbb{N}$ . Then any  $S$  from which exists a surjective map to  $T_k$  must have at least  $n$  members and  $S$  such that there is a bijection from  $S$  to  $T_n$  must have exactly  $n$  members.

## Step 2

2 of 4

Let us now examine  $T_{n+1} = \{t_1, t_2, \dots, t_n, t_{n+1}\}$  and its subset  $T_n = \{t_1, \dots, t_n\}$ . If  $\varphi$  is a surjection from  $S$  onto  $T_{n+1}$ , it is specially a surjection onto  $T_n$  as well. By the induction assumption,  $|S| \geq n$ . Assume that  $|S| = n$ . Then, again by the induction assumption,  $\varphi$  is a bijection from  $S$  to  $T_n$ . But then  $t_{n+1}$  is not in the image of  $\varphi$ , which is a contradiction with the surjectivity of  $\varphi$ .

We still need to show that  $|S| = n + 1$  implies that  $\varphi$  is a bijection. By the induction assumption, an  $n$ -element subset of  $S$  is bijective to an  $n$ -element subset of  $T_{n+1}$ , so let us denote the elements of  $S$  that map to  $(t_i)_{1 \leq i \leq n}$  as  $(s_i)_{1 \leq i \leq n}$  and the final element of  $S$  as  $s_{n+1}$ . Since  $\varphi$  is surjective, it must be mapped to  $t_{n+1}$ . Therefore,  $\varphi$  is bijection from  $\{s_{n+1}\}$  to  $\{t_{n+1}\}$ . Moreover, it is also a bijection from  $S = \{s_1, \dots, s_{n+1}\}$  to  $t = \{t_1, \dots, t_{n+1}\}$ , so the claim is proven by the principle of mathematical induction.

## Result

4 of 4

We prove this by induction on the amount of elements in  $|T|$ .

## 5. a

This is equivalent to the statement "If  $n$  is composite, then  $2^n - 1$  is composite.". Therefore, this is what we will prove. Let  $n$  be composite,  $n = ab$  with  $1 < a, b < n$ . If  $n$  is even, i.e.,  $n = 2k$  for some  $k \in \mathbb{N}$ , then  $2^n - 1 = 2^{2k} - 1 = (2^k - 1)(2^k + 1)$ . Notice that  $n \geq 4$ , which means that  $k \geq 2$  and neither  $(2^k - 1)$  nor  $(2^k + 1)$  can be equal to 1. Therefore,  $2^n - 1$  is then also composite.

Now, let us assume  $n$  is odd and  $n = ab$ . Then, since  $n$  is composite by assumption,  $n \geq 9$  and  $a, b \geq 3$ . Also, notice that  $a$  and  $b$  must also be odd. Now we have that:

$$2^n - 1 = 2^{ab} - 1 = (2^a)^b - 1^b = (2^a - 1) \left( (2^a)^{b-1} + (2^a)^{b-2} + \dots + 2^a + 1 \right),$$

where the last equality follows from a well-known and easily verifiable identity about the difference of odd powers. This indeed means that  $2^n - 1$  is composite because both factors are bigger than one by discussion above.

## Result

2 of 2

The proof techniques we use here are the contrapositive and splitting into cases, the only technical tools are formulas for the difference of odd and even powers.

## 6. a



We will denote the claim  $a_n = a_0 a_1 \dots a_{n-1} + 2$  by  $A_n$  and prove it by mathematical induction.

Firstly, let  $n = 1$ . Then:  $a_n = 2^{2^1} + 1 = 4 + 1 = 3 + 2 = 2^{2^0} + 2 = a_0 + 2$ , so the basis holds.

Assume that  $A_k$  holds for some  $k \in \mathbb{N}$  and let us consider  $A_{k+1}$ . It holds that:

$$\begin{aligned}
 a_{k+1} &= 2^{2^{k+1}} + 1 \\
 &= 2^{2^k \cdot 2} + 1 \\
 &= 2^{2^k} \cdot 2^{2^k} + 1 \\
 &= (a_k - 1)(a_k - 1) + 1 \\
 &\stackrel{A_k}{=} (a_0 a_1 \dots a_{k-1} + 1)(a_k - 1) + 1 \\
 &= a_0 a_1 \dots a_{k-1} a_k - a_0 a_1 \dots a_{k-1} + a_k - 1 + 1 \\
 &\stackrel{A_k}{=} a_0 a_1 \dots a_{k-1} a_k - a_0 a_1 \dots a_{k-1} + (a_0 a_1 \dots a_{k-1} + 2) \\
 &= a_0 a_1 \dots a_{k-1} a_k + 2
 \end{aligned}$$

This is exactly  $A_{k+1}$  so the claim is proven by the principle of mathematical induction.

## Result

2

This is again a standard example of a proof by induction requiring just a bit of algebraic agility.

## 7. a

A non-constant polynomial that cannot be factored into the product of two non-constant polynomials is called irreducible polynomial.

To show that every polynomial with rational coefficients can be written as a product of irreducible polynomials;

This can be proved by using induction.

[Comment](#)

Step 2 of 4 ^

Let  $P(n)$  be every polynomial of degree  $n$ . This can be written as a product of irreducible polynomials.

For degree  $n = 1$ :

Every polynomial of degree 1 is irreducible, so it can be written as itself.

Assume that this statement is true for polynomials with degree  $1, 2, 3, \dots, n$ .

[Comment](#)

Step 4 of 4 ^

To show that it is true for polynomials with degree  $k = n + 1$ .

Every polynomial of degree  $k$  is either irreducible and then it could be written as a itself or it could be written as a product of two polynomials with smaller degree. But then use the assumption and write down each of them as a product of irreducible polynomials.

Therefore, every polynomial with rational coefficients can be written as a product of irreducible polynomials.

**Hence proved**

8. a

We will prove this claim using the induction axiom.

Let  $S = \{1\} \cup \{m' \mid m \in \mathbb{N}\}$ . Clearly,  $1 \in S$ , so condition (I) holds. Let  $n \in S$ . This implies that  $n \in \mathbb{N}$ , so  $n' \in S$  by definition of  $S$  and condition (II) holds as well. Now the induction axiom implies that  $S = \mathbb{N}$ .

## Result

2 of 2

Here we give what is really the quintessential induction proof since it follows directly from the induction axiom and describes the structure of the set of natural numbers.

9. a

(a) We will use associativity of addition that was proven in the chapter and denote it by (I). The outline of the proof is sort of a double induction: Let us denote the claim  $m + n = n + m$  by  $C_{n,m}$ . We will first do an induction on  $n$ . Note that throughout the exercise, if the argument for an equality is not given, it follows from the definition of the operations.

Let  $n = 1$ .  $C_{1,m}$  is  $m + 1 = 1 + m$ . We will prove  $C_{1,m}$  by induction on  $m$ . For  $m = 1$ ,  $C_{1,m}$  clearly holds. Assume that  $C_{1,j}$  holds for some natural  $j$ . Now, for  $j'$ , it holds that:

$$j' + 1 = (j + 1) + 1 \stackrel{C_{1,j}}{=} (1 + j) + 1 \stackrel{(I)}{=} 1 + j'.$$

This proves the "subinduction", so  $C_{1,m}$  holds for all natural  $m$ .

Now assume that  $C_{i,m}$  holds for all some  $i$  and all natural  $m$ .

Let us prove that then it also holds for  $i'$  and all natural  $m$ :

$$\begin{aligned} m + i' &= m + (i + 1) \\ &\stackrel{(I)}{=} (m + i) + 1 \\ &\stackrel{C_{i,m}}{=} (i + m) + 1 \\ &\stackrel{(I)}{=} i + (m + 1) \\ &\stackrel{C_{1,m}}{=} i + (1 + m) \\ &\stackrel{(I)}{=} (i + 1) + m \\ &= i' + m \end{aligned}$$

This concludes the proof, the claim follows from the principle of mathematical induction.

(b) We will use (c) to prove associativity so the reader is encouraged to read (c) first. We need to prove that the claim  $P_c : (ab)c = a(bc)$  holds for all natural numbers  $a, b$  and  $c$ . We will do this by induction on  $c$ .

For  $c = 1$ , it holds that  $(ab)1 = ab = a(b \cdot 1)$ , so the induction basis holds.

Assume now that the claim holds for  $c = n$ . Then, for  $n'$  it holds that:

$$(ab)n' = (ab)n + ab \stackrel{P_n}{=} a(bn) + ab \stackrel{(c)}{=} a(bn + b) = a(bn')$$

Therefore, the associative law holds by the principle of mathematical induction.

### Step 3

3 of 5

(c) We need to prove that the claim  $P_c : a(b + c) = ab + ac$  holds for all natural numbers  $a, b$  and  $c$ . We will do this by induction on  $c$ .

For  $c = 1$ , it holds that  $a(b + 1) = ab' = ab + a = ab + a \cdot 1$ , so the induction basis holds.

Assume now that the claim holds for  $c = n$ . Then, for  $n'$  it holds that:

$$a(b + n') = a(b + n)' = a(b + n) + a \stackrel{P_n}{=} ab + an + a = ab + an'$$

Therefore, the distributive law holds by the principle of mathematical induction.

(d) We need to prove that the claim  $P_c : (a + b = a + c \implies b = c)$  holds for all natural numbers  $a, b$  and  $c$ . We will do this by induction on  $c$ .

For  $c = 1$ , assume that  $b \neq 1$ . Then there exists a  $d \in \mathbb{N}$  such that  $b = d'$ . Now let  $a + b = a + 1$ . This is equivalent to  $a' = a + 1 = a + b = a + d' = (a + d)'$ , but this is a contradiction with the injectivity of the successor function, so the induction basis holds.

Assume now that  $P_c$  holds for  $c = n$ . Then, for  $n'$ , let  $a + b = a + n'$ . By the exact same argument as in the proof of the induction basis,  $b$  cannot be equal to 1. Therefore, let  $b = d'$ . Now we have that  $a + d' = a + n'$ , which is by definition equivalent to  $(a + d)' = (a + n)'$ . By the injectivity of the successor function, it holds that now  $a + d = a + n$ . By the induction assumption  $P_n$ , then  $d = n$  and so  $b = n'$ .

Therefore, the cancellation law holds by the principle of mathematical induction.

### Result

5 of 5

To obtain the solution to this exercise, we use the Peano axioms, most notably induction in all subexercises. The trickiest one is the "double induction" in (a).

## 10. a

(a) If  $a < b$ , then  $b = a + m$  for some  $m \in \mathbb{N}$ . Then it holds that  $b + n = (a + m) + n = a + (m + n) = a + (n + m) = (a + n) + m$ .

Therefore, by commutativity and associativity of the natural numbers,  $b + n = (a + n) + m$ , so  $a + n < b + n$ .

### Step 2

2 of 4

(b) Let  $a < b$  and  $b < c$ . Then it holds that  $b = a + n$  and  $c = b + m$  for some  $n, m \in \mathbb{N}$ . This implies that  $c = b + m = (a + n) + m = a + (n + m)$ . Thus,  $a < c$ , so the relation  $<$  is indeed transitive.

(c) Let us prove this by induction on  $a$ . Firstly, let  $a = 1$ . If  $b = 1$ , then  $a = b$ . Now assume  $b \neq 1$ . Then there exists a natural number  $c$  such that  $b = c'$ . In other words,  $c = b + 1 = 1 + b$ . Therefore, by definition,  $1 < c < c' = b$ . Let the claim hold for some  $a = k \in \mathbb{N}$ . Let us now examine  $k'$ . By induction assumption, for all natural  $m$  it holds that  $m < k$ ,  $m = k$  or  $k < m$ . If  $m < k$ , then  $k = m + l$  for some  $l \in \mathbb{N}$ . But then  $k' = k + 1 = (m + l) + 1 = m + (l + 1)$ , so  $m < k'$  as well.

Now consider  $m = k$ . Then clearly  $k' = k + 1 = m + 1$ , so  $m < k'$ .

Finally, let  $k < m$ . Then  $m = k + d$ , for some natural number  $d$ . If  $d = 1$ , then  $k' = k + 1 = m$ , so  $k' = m$ . On the other hand, if  $1 < d$ , then  $k' = k + 1 \stackrel{(a)}{<} k + d = m$ , so  $k' < m$ .

We exhausted all the options and we acquired that the only possibilities are  $m < k'$ ,  $m = k'$  or  $k' < m$ , hence the claim follows from the principle of mathematical induction.

## Result

4 of 4

Here we use previous findings about commutativity and associativity to acquire some properties of the  $<$  relation.

## 11. a

Assume that the set  $S$  has the property listed in the exercise, but it is not equal to  $\mathbb{N}$ . Then  $T := \mathbb{N} \setminus S$  is non-empty. Thus, it has a minimal element  $a$ . Since  $a$  is the minimal element of  $T$ , all  $n \in \mathbb{N}$  such that  $n < a$  are also in  $S$ . But then the property of  $S$  implies that  $a \in S$  as well. This is a contradiction, so the principle of complete (also called strong) mathematical induction holds.

## Result

2 of 2

Here we give a simple proof relying on a minimal element.

## 12. a

(a) Let  $S$  be a partially ordered set that contains an upper bound  $b$ . Assume that there exists another upper bound  $b'$ . By definition of upper bounds, it holds that

$$\begin{aligned}\forall s \in S, b &\geq s \\ \forall s \in S, b' &\geq s\end{aligned}$$

Specifically,  $b \geq b'$  and  $b' \geq b$ . By A.3.1 (iii), also called the antisymmetric property, it holds that  $b = b'$ . Therefore, the upper bound in a partially ordered set is unique.

## Step 2

2 of 4

Assume for the sake of contradiction that  $b$  is not a maximal element, i.e., it is false that it does not exist an element  $a$  of  $S$  such that  $b \leq a$ . Therefore, there exists an  $a$  such that  $b \leq a$ . By definition of the upper bound,  $a \leq b$ . Therefore, again by the antisymmetric property,  $a = b$ , so  $b$  is the maximal element of  $S$ .

## Step 3

3 of 4

(b) Let  $m$  be a maximal element, that is, such that there does not exist  $a \in S$  such that  $m < a$ . Since  $S$  is totally ordered, it holds that  $m \geq a$  for all  $a \in S$ . This is exactly the definition of an upper bound, so the claim is proven.



## Result

4 of 4

In this exercise, we practice getting comfortable with the definitions of maximal elements in upper bounds in order to be prepared for the setting Zorn's Lemma is used in.

13. a

Let  $I$  be an ideal in  $R$  different from  $R$  itself. Let  $S$  be the set of all ideals  $J$  such that  $I \subseteq J$  and  $J \neq R$ .  $S$  is now a partially ordered set with respect to  $\subseteq$ .  $I$  is then contained in  $S$ , so  $S$  is non-empty. Let  $T$  be a totally ordered subset of  $S$ . We want to show that  $T$  has an upper bound. Let us define  $K := \bigcup_{J \in T} J$ . Since every set in  $S$  contains  $I$ , also every set in  $T$  contains  $I$ , so  $K$  contains  $I$ . A union of ideals is an ideal, so  $K$  is an ideal. Also, we claim that  $K \neq R$ .

Assume for the sake of contradiction that  $K = R$ , then  $1 \in K$ , so  $1 \in J$  for some  $J \in S$ , which implies  $J = R$ . On the other hand  $R \notin S$  by definition, so we've arrived at a contradiction and it holds that  $K \neq R$ .

Because  $K$  is an ideal,  $K \supseteq I$  and  $K \neq R$ , it holds that  $K \in S$ .  $K$  is clearly an upper bound of  $T$  by definition, which means that  $S$  is inductive. Zorn's lemma now implies that  $S$  has a maximal element, so every ideal  $I \neq R$  of a ring  $R$  is contained in a maximal ideal.

## Result

2 of 2

This is a standard Zorn's lemma proof where we focus on proving the inductivity of the set we are considering.

14. a

(a)

To show that:

$$\frac{\partial f}{\partial y} = \frac{\partial f_0}{\partial y_0} + \frac{\partial f_1}{\partial y_0} i$$

[Comment](#)

Step 2 of 5 ^

Let the polynomial having complex coefficients  $(f_0, f_1)$  be a function of  $(y, y_0)$ , where  $f$  is function of  $(x, y)$ .

$$f(y_0) = f_0(y_0) + if_1(y_0)$$

The above equation is in the form of  $z = a + ib$  which is a complex number.

Partially differentiating the equation with respect to ' $y$ ',

$$\frac{\partial(f_0)}{\partial y} = \frac{\partial(f_0 y_0)}{\partial y_0} + i \frac{\partial(f_1 y_0)}{\partial y_0}$$

This implies that,

$$\frac{\partial f}{\partial y} = \frac{\partial f_0}{\partial y_0} + i \frac{\partial f_1}{\partial y_0}$$

This implies that,

$$\frac{\partial f}{\partial y} = \frac{\partial f_0}{\partial y_0} + \frac{\partial f_1}{\partial y_0} i$$

Hence proved

(b)

To show that:

$$\frac{\partial f_0}{\partial y_0} = \frac{\partial f_1}{\partial y_1}; -\frac{\partial f_0}{\partial y_1} = \frac{\partial f_1}{\partial y_0}$$

The above equations are called as Cauchy Riemann Equations.

In the field of complex analysis in mathematics, the Cauchy–Riemann equations consist of a system of two partial differential equations.

Let,

$$f(y_0, y_1) \equiv f_0(y_0, y_1) + if_1(y_0, y_1)$$

Where,

$$z = y_0 + iy_1$$

So, on differentiating both sides:

$$dz = dy_0 + idy_1$$

The total derivative of  $f$  with respect to  $z$  is then,

$$\begin{aligned} \frac{df}{dz} &= \frac{\partial f}{\partial y_0} \frac{\partial y_0}{\partial z} + \frac{\partial f}{\partial y_1} \frac{\partial y_1}{\partial z} \\ &= \frac{1}{2} \left( \frac{\partial f}{\partial y_0} - i \frac{\partial f}{\partial y_1} \right) \end{aligned}$$

Now use  $f(y_0, y_1) \equiv f_0(y_0, y_1) + if_1(y_0, y_1)$ ;

$$\frac{df}{dz} = \frac{1}{2} \left[ \left( \frac{\partial f_0}{\partial y_0} + i \frac{\partial f_1}{\partial y_0} \right) - i \left( \frac{\partial f_0}{\partial y_1} + i \frac{\partial f_1}{\partial y_1} \right) \right]$$

Along the  $y_0$ -axis,  $\frac{\partial f}{\partial y_1} = 0$ , to get

$$\frac{df}{dz} = \frac{1}{2} \left( \frac{\partial f_0}{\partial y_0} + i \frac{\partial f_1}{\partial y_0} \right) \dots\dots (1)$$

Along the  $y_1$ -axis,  $\frac{\partial f}{\partial y_0} = 0$ , to get

$$\frac{df}{dz} = \frac{1}{2} \left( -i \frac{\partial f_0}{\partial y_1} + \frac{\partial f_1}{\partial y_1} \right) \dots\dots (2)$$

If  $f$  is complex differentiable, then the value of the derivative must be the same for a given  $dz$ , regardless of its orientation.

From equation (1) and (2), to get

$$\begin{aligned} \frac{1}{2} \left( \frac{\partial f_0}{\partial y_0} + i \frac{\partial f_1}{\partial y_0} \right) &= \frac{1}{2} \left( -i \frac{\partial f_0}{\partial y_1} + \frac{\partial f_1}{\partial y_1} \right) \\ \frac{\partial f_0}{\partial y_0} + i \frac{\partial f_1}{\partial y_0} &= -i \frac{\partial f_0}{\partial y_1} + \frac{\partial f_1}{\partial y_1} \end{aligned}$$

Compare real and imaginary part, to get

$$\boxed{\frac{\partial f_0}{\partial y_0} = \frac{\partial f_1}{\partial y_1}}$$

And,

$$\boxed{-\frac{\partial f_0}{\partial y_1} = \frac{\partial f_1}{\partial y_0}}$$

Hence proved

15. a



A Surface  $S$  in  $\mathbb{R}^{n+1}$  is a set of the form  $g^{-1}(c)$  where  $g: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  is a smooth function and  $c \in \mathbb{R}$  along with the fact  $\nabla g(p) \neq 0$  where,  $\nabla g$  denotes the gradient of  $g$  and  $p \in S$  is any point of  $S$ . every surface is a 2 manifold.

[Comment](#)

Step 2 of 5 ^

To show that locus  $f = 0$  forms a manifold of dimension 2.

Consider the complex polynomial:

$$f(x, y)$$

Then,

$$f(x, y) = f_0 + f_1 i$$

Where,  $f_i = f_i(x_0, x_1, y_0, y_1)$  is a real valued function of four variables are  $x_0, x_1, y_0, y_1$ .

Such that,

$$x = x_0 + x_1 i$$

$$y = y_0 + y_1 i$$

Since,  $f$  is a polynomial in  $x$  and  $y$ .

So,  $f_i$  are real valued polynomials in the real variables  $x_i$  and  $y_i$ .

So, they have continuous partial derivatives.

Now, since partial derivatives of  $f_i$  are continuous and  $f_i$  being a polynomial function is continuous.

Hence,  $f(x, y)$  can treat separately in terms of real functions and is a smooth function.

Let,  $(a, b) \in \mathbb{C}^2$

Such that,

$$f(a, b) = 0$$

Where,

$a, b$  are arbitrary complex numbers.

Since,  $f = 0, \frac{\partial f}{\partial y} = 0, \frac{\partial f}{\partial x} = 0$  have no common solutions.

Then, either  $\frac{\partial f}{\partial y}(a, b) \neq 0$

Or,

$$\frac{\partial f}{\partial x}(a, b) \neq 0.$$

Without loss of generality assume that  $\frac{\partial f}{\partial y}(a, b) \neq 0$

So, by Implicit Function Theorem,

There exists an open neighborhood  $U$  of  $x$  in  $\mathbb{C}$  on which a unique continuous function  $Y(x)$  exists having properties:

$$f(x, Y(x)) = 0$$

And

$$Y(a) = b$$

Thus,

$$(x, Y(x)) = f^{-1}(0)$$

Let,

$$S = f^{-1}(0)$$

Also,

$$\nabla f(a,b) = \left( \frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b) \right)$$

Since,

$$\frac{\partial f}{\partial y}(a,b) \neq 0$$

Hence,

$$\nabla f(a,b) \neq 0 \text{ for all } (a,b) \in S$$

Thus,  $S$  is a surface.

Now, clearly locus  $f = 0$  forms the set  $S$  which is a surface.

Since, every surface is a 2 manifold.

Thus, locus  $f = 0$  forms a manifold of dimension 2.

**Hence proved**